Using Meta-Knowledge Mined from Identifiers to Improve Intent Recognition in Conversational Systems

Claudio S. Pinhanez, Paulo Cavalin, Victor Ribeiro, Ana Paula Appel, Heloisa Candello, Julio Nogima, Mauro Pichiliani, Melina Guerra, Maira Gatti de Bayser, Gabriel Malfatti, Henrique Ferreira

IBM Research - Brazil
csantosp@br.ibm.com

Abstract

In this paper we explore the improvement of intent recognition in conversational systems by the use of meta-knowledge embedded in intent identifiers. Developers often include such knowledge, structure as taxonomies, in the documentation of chatbots. By using neuro-symbolic algorithms to incorporate those taxonomies into embeddings of the output space, we were able to improve accuracy in intent recognition. In datasets with intents and example utterances from 200 professional chatbots, we saw decreases in the equal error rate (EER) in more than 40% of the chatbots in comparison to the baseline of the same algorithm without the meta-knowledge. The meta-knowledge proved also to be effective in detecting out-of-scope utterances, improving the false acceptance rate (FAR) in two thirds of the chatbots, with decreases of 0.05 or more in FAR in almost 40% of the chatbots. When considering only the well-developed workspaces with a high level use of taxonomies, FAR decreased more than 0.05 in 77% of them, and more than 0.1 in 39% of the chatbots.

1 Introduction

Classification of sentences into a discrete set of classes is a key part of professional conversational systems. In fact, most of those systems require developers to define the different classes, or *intents*, by enumerating exemplars of each of them, since classification is often performed using *machine learning (ML)* methods. The process of classifying an input sentence into a specific intent or signaling it as *out-of-scope (OOS)* of the system is often referred to as *intent recognition*.

Determining a class solely on a list of exemplars is a practical method to implement ML systems but it is hardly a natural way for human beings to define a class. In real life, people define classes often using a rich mix of symbolic definitions, sometimes taxonomic in nature, such as in "a credit card is a type of bank card", coupled with its sub-classes, for instance, "basic", "premium", and typical features such as "international". People also use exemplars, "card X of bank Y is a credit card", as well as particular examples to describe a sub-class, such as in "card W is an international card". They also use counter examples, either categorically, "a debit card is not a credit card", or in examples, "card Z is not a credit card". Defining and specifying classes in the real world is, in fact, a cultural, contextual, and linguistic construct, and how people and societies perform this process is a traditional research subject in social sciences, notably in anthropology (Durkheim and Mauss, 1963; Needham, 1979; Bowker and Star, 2000).

This paper explores algorithms for intent recognition which use both the sets of exemplars and taxonomic-like symbolic descriptions of a class to define and train intents in conversational systems using ML methods. We aim not only to provide methods more aligned to everyday class definition practices of developers but also to improve the accuracy of the ML methods. Inspired by a reverse dictionary algorithm (Kartsaklis et al., 2018) and previous work on keyword-based classification (Cavalin et al., 2020), we propose three *neuro-symbolic* algorithms which combine taxonomic descriptions of classes with traditional exemplar-based supervised learning. We show that those novel algorithms are able to decrease error rates for a significant number of datasets, particularly in the difficult task of detecting OOS cases in real, professional chatbots.

The key idea behind our algorithms is to substitute the typical *softmax* used in the output layer of a ML text classifier with a space of embeddings of the taxonomic descriptions of the intents. The training process uses the exemplars in standard ways while the recognition process is performed using similarity distances in the embedding output space.

Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, pages 7014–7027
August 1–6, 2021.

This is similar to ideas used in *zero-shot learning* methods (Palatucci et al., 2009; Socher et al., 2013; Akata et al., 2015, 2016), in which classes defined by sub-concepts are also encoded with special embeddings to allow detection of new classes without exemplars.

We tested our algorithms using real datasets by exploring a common practice among developers of conversational systems, who often embed symbolic knowledge as documentation in intent identifiers. In a previous work (Pinhanez et al., 2021), we observed a pattern among developers of using taxonomic-like structures to name the intents in which strings of reoccurring concepts are used to identify and document the different classes. For example, an intent about utterances where users ask for the balance of a credit card may be named "checking_credit_card_balance", while an intent related to finding out the date of payment of the balance could be identified as "asking_credit_card_balance_payment_date".

We call those structures *intent proto-taxonomies*, and real examples are shown in figures 1 and 2. In (Pinhanez et al., 2021), we studied the use by developers of intent proto-taxonomies quantitatively and qualitatively, as well as proposed an algorithm to mine this meta-knowledge automatically, and concluded that their use is fairly common in at least one professional chatbot development platform. This paper focuses on the algorithms to use the meta-knowledge and on evaluating their impact on the accuracy of intent recognition.

The paper starts by looking into the recent advances in neuro-symbolic systems and describing briefly the practice of developers of conversational systems of embedding meta-knowledge within the source code of their systems. We follow by describing the proposed three algorithms integrating such meta-knowledge into intent recognition ML algorithms and by evaluating them first with two typical intent recognition datasets, and then with hundreds of workspaces created in a professional tool called here *ChatWorks*[1]. The results show most of those workspaces can benefit from the techniques described in this paper, notably for OOS detection tasks, often with accuracy improvements of 5% or more solely derived from the use of the additional symbolic description from the documentation.

[1]We use an anonymous name for the tool due to publication restrictions from the platform company.

2 Related Work

The value and limits of symbolic categorization in AI have been of interest since the early days (Newell, 1973; Richards, 1982; Kosslyn, 2006). But our work fits more in the context of a growing belief that symbolic knowledge needs to be included in ML systems, materialized in the so called *neuro-symbolic* approaches (Parisotto et al., 2017; Besold et al., 2017; Tenenbaum et al., 2011; Bengio, 2017; Mao et al., 2019; Hudson and Manning, 2019a; De Raedt et al., 2019).

Neuro-symbolic methods *"aim to transfer principles and mechanisms between (often nonclassical) logic-based computation and neural computation"* (Besold et al., 2017). Such kind of systems are viewed by some researchers as a way to embed high-level knowledge and even some form of "consciousness" into machine learning systems, making the language to develop them closer to *"what passes in a man's own mind"* (Bengio, 2017).

In recent years, AI has witnessed a myriad of novel neuro-symbolic techniques and their application to different problems, contexts, and scenarios (Parisotto et al., 2017; Manhaeve et al., 2018; d'Avila Garcez et al., 2019; Hudson and Manning, 2019b; De Raedt et al., 2019). For instance, in (Mao et al., 2019), an approach for image understanding is suggested which takes the object-based scene representations and translates sentences into executable, symbolic programs. In (Oltramari et al., 2020), embeddings of knowledge graphs are used as attention layers for tasks such as autonomous driving (AV) and question-answering. And in (Kartsaklis et al., 2018), random walks in a knowledge graph are mapped as sentence embeddings for use in an *inverse dictionary problem*.

One important requirement for many neuro-symbolic systems is to represent knowledge in a structured format such as knowledge graphs, ontologies, or taxonomies (Ji et al., 2020). In some cases, such as the scene ontology for autonomous vehicles in (Oltramari et al., 2020), a lot of effort was needed for manual annotation. Nevertheless, as presented in (Fossati et al., 2015), an unsupervised approach can sometimes be used to mine the meta-knowledge introduced by the experts, such as the classes in *Wikipedia* pages.

Considering our context, intent identifiers are sometimes described using high-level representations of the class as we detail later. This is similar to what is used in some *zero-shot learning*

techniques (Wang et al., 2019) in which classes manually defined by sub-concepts are encoded with special embeddings so new classes can be detected without training (Palatucci et al., 2009; Socher et al., 2013; Akata et al., 2015, 2016; Chen et al., 2016). In (Chen et al., 2016), for example, intent identifiers can be formatted as natural language sentences to learn a model which maps training examples into those sentences, so that the meta-knowledge can be used in zero-shot learning. However, the dataset explored in that work is very limited. Recent work has also demonstrated that intent recognition can be improved by enhancing class representations with "keywords" which are extracted from exemplar utterances considering their most common words (Cavalin et al., 2020).

This work focuses on high level class representations based on taxonomies and aims to explore their usefulness as enhancers of ML intent recognition algorithms. It also explores different ways of embedding taxonomy-like meta-knowledge considering different methods of representation.

3 The Knowledge Embedded in Intent Identifiers of Conversational Systems

Most real-world, deployed conversational systems in use today have been built based on the rule-based *intent-action* paradigm, using platforms such as *Luis.ai*, *Watson Assistant*, or *Alexa Skills*. Each *intent* corresponds to a desired information or answer from the user and is defined by a set of exemplar utterances by the chatbot developers. During runtime, each utterance from the user is recognized as one of the defined intents or as *out-of-scope* (OOS), and then the associated *action* is outputted.

In the context of the chatbots built using the ChatWorks platform explored in this paper, a previous work of the authors of this paper (Pinhanez et al., 2021) has shown that the curators and developers of chatbots often store symbolic knowledge in a taxonomic form about the intent classes in a documentation field called *nameId*. Figure 1 shows some examples of those nameIds, obtained from a professional finance chatbot, here translated from the original in Portuguese and anonymized to preserve confidential information.

This practice was studied in workshops with developers (Pinhanez et al., 2021), which determined that the goal of the taxonomic description is to provide the intent classes with a summarized description of each intent. Such taxonomic naming

intent	INTENT ID (translated and anonymized)
1	inv_XXX-fidelity-concept-special_inv-kareen_27-08-YY
2	inv_XXX-general-check-balance
3	inv_XXX-general-check-investment_period
4	inv_XXX-general-check-issue_time-special_inv
5	inv_XXX-general-check-rate-management
6	inv_XXX-general-concept-special_inv
7	inv_XXX-general-desambiguation-check_withdraw_date--
8	inv_XXX-general-desambiguation-modalities-special_inv
9	inv_XXX-general-desambiguation-withdraw--
10	inv_XXX-general-income_tax-bank_website
11	inv_XXX-general-liquidity-special_inv
12	inv_XXX-general-withdraw-minimum_time-special_inv
13	inv_XXX-general-withdraw-using_selfservice-management
14	inv_XXX-easy-concept-special_inv-kareen_27-08-YY
15	inv_XXX-prefix-concept-special_inv-kareen_27-08-YY
16	inv_XXX-prefix-investment-minimum_time-bank_website
17	inv_XXX-prefix-withdraw-procedure-bank_website
18	inv_XXX-withdraw-phone_app--
19	inv_XXX-withdraw-using_app_bank-management
20	inv_XXX-withdraw-using_internet_banking--

Figure 1: Some nameIds of intents of a finance chatbot.

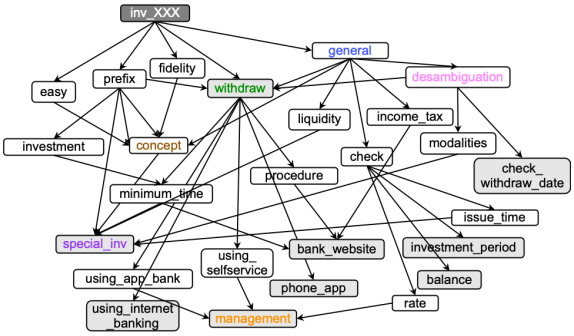

Figure 2: The intent proto-taxonomy associated to the intents of fig. 1.

patterns are also common in the way people organize files and e-mails in computers (Civan et al., 2008; Whittaker et al., 2011) and how software developers name functions (Yang et al., 2019).

As described in (Pinhanez et al., 2021), such knowledge-embedding practices are, in fact, fairly common among curators in the ChatWorks platform. Using the algorithm reproduced in appendix A, taxonomic-like symbolic knowledge was automatically extracted from *workspaces* defining almost 7,000 professional chatbots, in two different languages. By considering the different words in the nameIds as basic concepts and consecutive concepts as having connections between them, we can structure the set of nameIds as a very basic *knowledge graph* (Ehrlinger and Wöß, 2016), hereby referred as an *intent proto-taxonomy*.

Figure 2 depicts the intent proto-taxonomy associated to the nameIds in fig. 1. Next, as proposed

in (Pinhanez et al., 2021), it is possible to compute the *taxonomy rate* of a workspace by calculating the ratio between the number of intents with taxonomies and the total number of intents.

In 3,840 professsional workspaces in the English language, it was found that 76% of them had a taxonomy rate above 10%, almost 52% had a taxonomy rate above 50%, and 16% had a very high taxonomy rate, above 90%. Moreover, the distribution followed a sort of "step" function where, as the threshold of 32 in the number of intents in a workspace was crossed, the majority of the workspaces had a taxonomy rate of more than 50%. It seems that, as the complexity of the workspace increases with the number of intents, more often developers resort to document them using an intent proto-taxonomy (see appendix B for details).

The use in our work of the intent proto-taxonomies as a symbolic description of classes is feasible because: (1) they are part of the documentation of the conversational system, so there is no need of acquiring knowledge from experts; (2) they are easily mined, as described in appendix A.

4 Using Taxonomic Intent Descriptions to Improve Intent Recognition

We present now a formal description of the methodology employed in this work which takes advantage of the intent proto-taxonomies using a neuro-symbolic approach. It expands some previous work which focused on the use of keywords as the source of symbolic information (Cavalin et al., 2020). .

4.1 Embedding the Set of Classes

An *intent classification* method is a function D which maps a set of sentences (potentially infinite) $S = \{s_1, s_2, ...\}$ into a finite set of classes $\Omega = \{\omega_1, \omega_2, ..., \omega_c\}$:

$$D : S \rightarrow \Omega \quad D(s) = \omega_i \tag{1}$$

To enable a numeric, easier handling of the input text, an embedding $\xi : S \rightarrow \mathbb{R}^n$ is often used, mapping the space of sentences S into a vector space $\mathbb{R}^n$, and defining a classification function $E : \mathbb{R}^n \rightarrow \Omega$ such that $D(s) = E(\xi(s))$. In most intent classifiers, E is composed of a function M which computes the likelihood of s being in a given class, often a neural network, followed by some sort of $argmax$ function. Typically, $softmax + argmax$ is used, noted simply as $softmax$ here:

$$S \xrightarrow{\xi} \mathbb{R}^n \xrightarrow{M} \mathbb{R}^c \xrightarrow{softmax} \Omega \tag{2}$$

This paper explores how to use embeddings in the output side of the classification function, that is, by embedding the set Ω of classes into another vector space $\mathbb{R}^m$, in some ways resembling the combination of object-based recognition and symbolic programming in (Mao et al., 2019). Instead, we combine here standard intent recognition methods with an encoding of taxonomies in knowledge graph-like structures. The idea is to use class embedding functions which somehow capture the knowledge in the intent proto-taxonomies.

Formally, we use a *class embedding* function $\psi : \Omega \rightarrow \mathbb{R}^m$, its inverse ψ^{-1}, and a function $M : \mathbb{R}^n \rightarrow \mathbb{R}^m$ to map the two vector spaces so $D(s) = \psi^{-1}(M(\xi(s)))$.

$$S \xrightarrow{\xi} \mathbb{R}^n \xrightarrow{M} \mathbb{R}^m \xrightarrow{\psi^{-1}} \Omega \tag{3}$$

In this paper we explore three sentence embedding methods to implement ξ. We use a two-layer neural network as M and employ the standard *Mean Square Error* (MSE) as the inverse ψ^{-1}, to determine the closest embedding of each class $\omega_i \in \Omega$ to the output of M.

4.2 Adapting *Kartsaklis* Method (LSTM)

Our basic inspiration for the algorithms of this paper is a text classification method proposed in (Kartsaklis et al., 2018) for the inverse dictionary problem where text definitions of terms are mapped to the term they define. The embedding of the class set into the continuous vector space (equivalent to the ψ function in equation 3) is done by expanding the knowledge graph of the dictionary words with nodes corresponding to words related to those terms. Next, random walks are perfomed on the graph to compute graph embeddings related to each dictionary node, using the *DeepWalk* algorithm (Perozzi et al., 2014).

A *Long Short-Term Memory* (LSTM) neural network, composed of two layers and an attention mechanism, is used in (Kartsaklis et al., 2018) for mapping the input texts to the input embedding vector space. To map the two continuous vector spaces representing the definitions and the dictionary terms, a two-layer neural network M, learned from the training dataset, is used.

For this work, the knowledge graph is replaced by an *intent proto-taxonomy* G which associates each class to a node and connects to them nodes which correspond to meta-knowledge concepts related to the class. To better capture the sequential

aspect of the intent proto-taxonomies, we also connect each class node to bigrams of concepts, i.e., the concatenation of two subsequent concepts. We represent this by the function ζ, such as $\zeta(\Omega) = G$, which is invertible. Substituting this in equation 3,

$$S \overset{LSTM}{\rightarrow} \mathbb{R}^n \overset{M}{\rightarrow} \mathbb{R}^m \overset{DeepWalk^{-1}}{\rightarrow} G \overset{\zeta^{-1}}{\rightarrow} \Omega \quad (4)$$

In practice, we compute the mapping from the class embedding space into the class set, called here $InvG : \mathbb{R}^m \rightarrow \Omega$, simply by determining the distance d between the output point in $\mathbb{R}^m$ and the inverted projection of each class from Ω and then considering the closest class. That is, for each $w_i \in \Omega$, we consider the associated node in G and compute the mapping in $\mathbb{R}^m$ of that node:

$$InvG(x) = \arg\min_{w_i} d(x, DeepWalk(G(w_i)) \quad (5)$$

By substituting this function into equation 4, we obtain the algorithm we call here *LSTM+T*:

$$S \overset{LSTM}{\rightarrow} \mathbb{R}^n \overset{M}{\rightarrow} \mathbb{R}^m \overset{InvG}{\rightarrow} \Omega \quad (6)$$

For comparison, the traditional corresponding classification method is tested, where the graph embedding and associated functions are replaced by *softmax+argmax*. We call this *LSTM*:

$$S \overset{LSTM}{\rightarrow} \mathbb{R}^n \overset{M}{\rightarrow} \mathbb{R}^c \overset{softmax}{\rightarrow} \Omega \quad (7)$$

4.3 An Alternative to LSTM: USE

Recently, several new general-purpose language models that can be used for computing sentence embeddings have been proposed, among them the *Universal Sentence Enconder (USE)* (Cer et al., 2018). Such an approach consists of a *transformer* neural network (Vaswani et al., 2017), trained on varied sources of data, such as *Wikipedia*, web news, web question-answer pages and discussion forum. USE has achieved state-of-the-art results in various tasks, so we decided to try it in our experiments as an alternative to the LSTM for the embedding of input sentences.

In this work we employed the version 3 of the multilingual USE [2]. By replacing LSTM with USE in eq. 6 we obtain algorithm *USE+T*:

$$S \overset{USE}{\rightarrow} \mathbb{R}^n \overset{M}{\rightarrow} \mathbb{R}^m \overset{InvG}{\rightarrow} \Omega \quad (8)$$

[2]https://tfhub.dev/google/universal-sentence-encoder-multilingual/3

Like in the previous case, we also compute the USE algorithm with traditional discrete softmax outputs for comparison, called here simply *USE*:

$$S \overset{USE}{\rightarrow} \mathbb{R}^n \overset{M}{\rightarrow} \mathbb{R}^c \overset{softmax}{\rightarrow} \Omega \quad (9)$$

4.4 Alternatives to DeepWalk

To explore variants of algorithms for embedding the classes and also approaches which do not need to be trained from scratch and allow on-the-fly handling of meta-knowledge, we tried replacing DeepWalk with two different methods.

The first one consists of applying USE sentence embeddings also for the class embeddings, such as in eq. 10. To simplify notation, *emb* represents either LSTM or USE embeddings for the input text.

$$S \overset{emb}{\rightarrow} \mathbb{R}^n \overset{M}{\rightarrow} \mathbb{R}^m \overset{USE^{-1}}{\rightarrow} G \overset{\zeta^{-1}}{\rightarrow} \Omega \quad (10)$$

This approach is similar to the way DeepWalk works but instead of training the graph embeddings from scratch, the class embeddings are represented by the mean sentence embedding computed from different random walks starting in the class node. We name such methods *LSTM+S* and *USE+S*, for *emb* substituted by LSTM and USE, respectively.

Additionally, we also evaluate the replacement of DeepWalk by the *Convolutional Deep Structured Semantic Model* (CDSSM) proposed in (Chen et al., 2016), yielding the following algorithm where *emb* can be either LSTM or USE embeddings.

$$S \overset{emb}{\rightarrow} \mathbb{R}^n \overset{M}{\rightarrow} \mathbb{R}^m \overset{CDSSM^{-1}}{\rightarrow} G \overset{\zeta^{-1}}{\rightarrow} \Omega \quad (11)$$

The CDSSM model consists of a three-layer convolutional neural network trained for creating embeddings of intent identifiers represented as sentences. In this work, we input to CDSSM the sequence of concepts listed in the nameId of each intent. We refer to these algorithms as *LSTM+C* and *USE+C*, for *emb* being substituted with LSTM and USE, respectively.

An intuitive way to understand those methods is to consider *USE+T* using a taxonomy as if its concepts had just abstract meanings: only their relations matter. In comparison, *USE+S* considers the meaning of the concepts besides their relations, while *USE+C* regards each nameId as a sentence, almost as if the developer had inputted a written description of the intent.

4.5 Out-of-Scope Sample Detection

In this paper we are interested both in the problems of: (1) deciding whether an user utterance is *in-scope (IS)* or *out-of-scope (OOS)* of the system; and (2) determining to which class an IS utterance belongs. For the former, a rejection mechanism based on a pre-defined threshold is used since it can be easily applied to all of the methods described previously without the need neither for any specific training procedure nor OOS training data.

In detail, suppose that for each class $\omega_i \in \Omega$ there is a score denoted $\phi_i \in Z$, where $|Z| = |\Omega|$. Given that $\max(Z)$ represents the highest score associated to a class and that a rejection threshold θ has been defined on a validation set, samples can be classified as OOS whenever $\max(Z) < \theta$. If so, they are simply rejected, i.e., no classification output is produced for them. Otherwise, the sample is considered as in-scope and the classification is conducted normally.

The scores in Z are represented either by the *softmax* probability for the traditional softmax-based methods or by the similarity of sentence and intent embeddings for the proposed three approaches. For the latter, the similarity is computed by means of the dot product between the two embeddings.

5 Metrics, Datasets, and Experiments

In this section we present the experiments to evaluate the three algorithms described in the previous section, using each of the input embeddings LSTM and USE. We explore the impact on intent recognition both in terms of classifying correctly utterances (IS accuracy) and of finding which utterances are not covered by the intents (OOS accuracy).

5.1 Evaluation metrics

We employ a commonly-used metric for OOS dectection, *equal error rate (EER)* (Tan et al., 2019), which corresponds to the classification error rate when the threshold θ is set to a value where *false acceptance rate* (FAR) and *false rejection rate* (FRR) are the closest. These two metrics are defined as:

$$FAR = \frac{\text{number of accepted OOS samples}}{\text{total of OOS samples}} \quad (12)$$

$$FRR = \frac{\text{number of rejected IS samples}}{\text{total of IS samples}} \quad (13)$$

In addition, *in-scope error rate (ISER)* is considered to report IS performance, i.e. the error rate considering only IS samples when θ is set to zero, similar to the *class error rate* in (Tan et al., 2019). This metric is important to evaluate whether the proposed classification methods are able to keep up with the performance of the baselines in the main classification task.

5.2 The Larson and Telco Datasets

During the development and initial testing of the algorithms, we used two English datasets for in-depth experimentation. The first is the publicly-available *Larson* dataset (Larson et al., 2019); the second is a private real-world chatbot dataset used by a telecommunications provider for customer care, called here the *Telco* dataset. In the Larson dataset, we created an intent proto-taxonomy by hand, expanding the original identifiers of intents. The goal of the adjustments was to avoid spurious interference from taxonomy shortcomings or errors in the results. The complete list of the created taxonomic description of intents is listed in the appendix C to allow the reproduction of our results and further experimentation. In the Telco dataset, we created by hand the intent proto-taxonomy.

In the Larson dataset there is a total of 22,500 IS exemplars, evenly distributed across 150 classes, where 18,000 were used for training and 4,500 for testing. We conducted a simulation of OOS detection with the IS exemplars by doing 5 random samplings where we took out 30 intents and 3,600 training exemplars. We trained only with the remaining 120 intents and 14,400 exemplars. The test was then conducted using all the non-used 4,500 exemplars, where the 3,600 associated to the trained classes were considered the IS samples and the remaining 900 became OOS samples.

The Telco dataset contains 4,093 exemplars and 87 intents. From those, 3,069 exemplars were used for training and 1,024 for testing. The OOS scenario was simulated by extracting different random samplings where 5 intents were removed. Given the smaller size of this dataset compared to Larson, we conducted 20 samplings instead of 5.

For both sets we considered the following setup defined after preliminary evaluations. For the LSTM-based methods, the input sentence embedding size was set to 150 and output embeddings to 200. DeepWalk walk sizes were set to 20 for LSTM+T and USE+T. For both USE- and softmax-based methods we trained a two-layer neural network with 800 hidden neurons for 50 epochs.

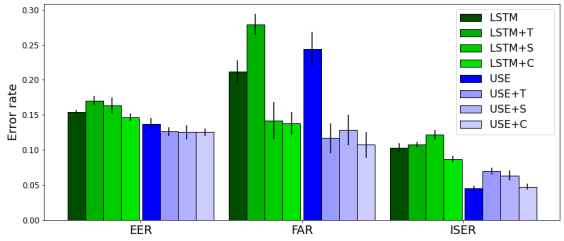

Figure 3: Different methods to incorporate the intent proto-taxonomy in Larson dataset, compared to the LSTM and USE baselines.

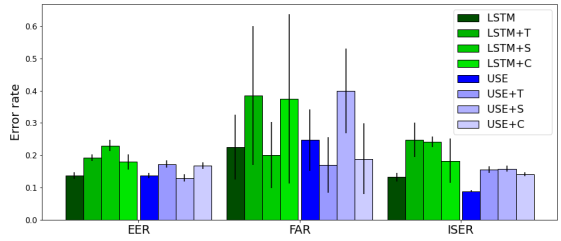

Figure 4: Different methods to incorporate the intent proto-taxonomy in Telco dataset, compared to the LSTM and USE baselines.

5.3 Results in the Larson and Telco Datasets

The results on the Larson dataset are graphically depicted in fig. 3. We observed that there was a slight improvement (a decrease) in EER, especially with the USE-based and the LSTM+C methods. More notably, there was a significant improvement in terms of FAR for all USE-based methods and LSTM+S and LSTM+C. Notice that even though the proposed approaches generally did not outperform LSTM and USE in ISER (except LSTM+C), we observed that the methods with better ISER tended to produce also better EER and FAR.

In fig. 4, we see that the results on the Telco dataset presented a different scenario. The proposed methods generally performed worse than or, at best, similar to LSTM and USE in EER. In terms of FAR, some methods such as USE+T and USE+C seem to outperform the others but, considering the high standard deviations, the results were not significant. On the other hand, we also observed that the methods failed to get close in ISER compared to the softmax-based methods. That seems to indicate that for the cases where making use of meta-knowledge harms too much ISER, the symbolic knowledge did not decrease neither EER nor FAR.

There were two key findings from our experiments with the Larson and the Telco datasets. First, the improvements using LSTM or USE as a baseline seemed to be similar, possibly slightly better for the USE algorithm. Second, and most importantly, we saw much more improvement in the use of the intent proto-taxonomy in the Larson than in the Telco dataset, in spite of the similar nature of the datasets and the intent proto-taxonomies. This motivated us to try out the ideas in a larger and more diverse number of workspaces and solely focusing on USE to simplify the experiments.

5.4 The ChatWorks Dataset

To test our algorithms in a context of high diversity and realism, we used the same large set of real, professional workspaces explored in (Pinhanez et al., 2021), which come from the professional chatbot development platform *ChatWorks*.

We started with the 3,840 workspaces available in English. To eliminate possible problems due to workspaces with poor quality, we employed the 3σ-rule, where values smaller greater than 3 standard deviations from the mean are not considered.Workspaces with the number of intents or exemplars below and above those thresholds were removed. Also, to avoid workspaces with few exemplars per intent, the ratio of the number of exemplars to the number of intents had to be greater than 10. From the filtered set we randomly selected 200 workspaces for testing.

The evaluation involved the execution of 20 iterations for each workspace. The tests were performed for all USE-based methods (USE, USE+T, USE+S, and USE+C). First, the workspaces were split into training and test datasets (75% and 25%, respectively). Next, the four methods were trained and tested on these datasets. The evaluation metrics (EER, FAR, and ISER) were then measured on the results for the test datasets and the average errors and their standard deviations were computed.

5.5 Results in the ChatWorks Dataset

Appendix D contains a table with the results for each of the 200 workspaces in the ChatWorks dataset. Figure 5 summarizes the results of the experiments showing the distribution of the 200 workspaces according to ranges of the improvement of each of the three methods compared to the baseline of USE. Improvement is calculated by subtracting the errors in each of our proposed methods from the errors in the USE baseline (error values are scores between 0 and 1). When one of our methods was worse than the baseline then $diff < 0$,

since smaller is better, and conversely for when it is better than the baseline, i.e., $diff \geqslant 0$.

The results shown in fig. 5 indicate that the USE+C algorithm achieved the best results in all three metrics, although there is a significant portion of workspaces where the other methods also did well, especially in OOS detection (FAR).

But, more important, the results seem to support our claim that meta-knowledge embedded in the output layer of our neuro-symbolic algorithms can improve intent recognition performance in practical systems. Notably in OOS detection (FAR), 67% of the workspaces experienced a decrease in the error rate using USE+C. Besides, in 39% of the workspaces we observed a decrease in the error rate of more than 0.05 (in a 0 to 1 scale), and in 23%, of more than 0.1. The USE+T also did well with similar but slightly smaller decreases in error.

Overall, the error rates for the EER metric also decreased in relation to the baseline. Figure 5 shows that 41% of the workspaces had some level of decrease in EER with the USE+C algorithm, in 10% of them with decreases of 0.05 or more. However, the results for the in-scope accuracy (ISER) were much smaller with only about 16% of the workspaces having any kind of decrease.

The ChatWorks dataset, as noted before, includes all kinds of workspaces. Taxonomy rates varies anywhere from 0 to 1, and there are very small and very large workspaces. To test our methods in a scenario closer to a professional, well-developed chatbot, we filtered further the dataset to include only workspaces with taxonomy rate greater or equal to 0.7, with number of intents equal or more than 32, and at least an average of 25 exemplars per intent, resulting in 18 workspaces.

Figure 6 shows the distribution of the results of the experiments with those 18 workspaces, which were better than in the full ChatWorks dataset. Both USE+T and USE+C yielded EER decreases in 50% or more of the workspaces. Moreover, 83% of the workspaces decreased the FAR error, either with USE+T or USE+C, and both decreased FAR in more than a third of the workspaces by more than 0.1. We discuss the results and implications next.

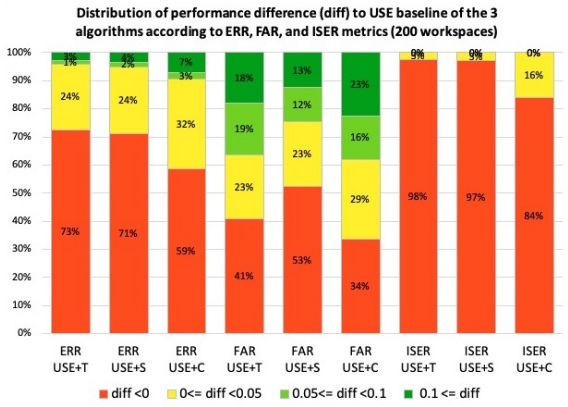

Figure 5: Distribution of performance difference (diff) to USE baseline of the 3 methods according to EER, FAR, and ISER metrics in all 200 workspaces of the Chatworks dataset.

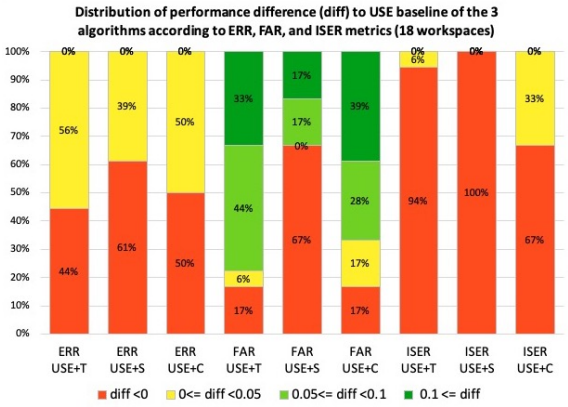

Figure 6: Distribution of performance difference (diff) to USE baseline of the 3 methods according to EER, FAR, and ISER metrics in the 18 most developed workspaces of the Chatworks dataset.

6 Discussion and Future Work

We started this paper by proposing the combination of exemplars and symbolic characterizations of a class as a way to enhance ML-based intent recognition. We proposed 3 new neuro-symbolic algorithms and tested them using datasets built using data from intent identifiers of conversational systems. Such identifiers often store taxonomic-like structures, due to a common practice among developers of professional conversational systems (Pinhanez et al., 2021). The results of the experiments indicate that the intent proto-taxonomies embedded by those developers can indeed be used by many workspaces to improve accuracy in intent recognition, notably in OOS detection.

We see as one of the main contributions of this paper the creation of methods with which ML engineers can improve the accuracy of their systems by simply mining "documentation" from chatbots, without any further data and annotation.

Our results show that almost 40% of the 200 professional workspaces drawn from ChatWorks saw decreases of more than 0.05 in OOS detection

error rates. Also, in 42% of them the overall error rate was improved, using the USE+C algorithm. When considering the more well-structured and developed 18 workspaces, we saw much higher gains with the USE+T algorithm. Those accuracy improvements were achieved without any change in the training set but simply by incorporating the meta-knowledge into intent recognition.

Notice that the testing methodology used in this work is considerably harder than the practice of the majority of research papers, since it evaluates performance in 200 professional, non-edited workspaces from different domains. In reality, most ML algorithms do not perform well in all datasets, and ML practitioners often test different algorithms and parameters until accuracy is good enough.

However, the improvement in OOS detection (FAR) was not mirrored in classification error (ISER). First, we must keep in mind that intent classification is often performed in two steps, first OOS sentences detection and removal, followed by intent classification of the IS sentences. Given the improvements observed in OOS detection, it would make sense to use our algorithms in the first step for many of the ChatWorks workspaces (about 60% of them), and selectively use it for IS classification only when it works better than the baseline.

But why were there so many workspaces where we did not see impact? It is important to take into account that the ChatWorks dataset has workspaces in different stages of development and deployment. By selecting better quality workspaces, we saw much higher gains. We explored briefly characterizations of the intent proto-taxonomy quality, such as taxonomy rate, depth of the taxonomy, and number of concepts, but we saw no clear correlation with decreases in error rates. We believe more complex metrics of knowledge structure need to be employed to characterize which intent proto-taxonomies are likely to have the greatest impacts. We plan to do so in our future work.

It is important to notice that, in the workspaces where we did see impact, the symbolic knowledge was mined from an absolutely "raw" format. In spite of that, by using the basic graph mining method described in appendix A, it was possible to obtain a "meaningful" taxonomic structure, similar to a knowledge graph which could be used by our neuro-symbolic algorithms. To improve the quality of the taxonomies, we are working on designing an interface which allows the developers to manipulate directly the intent proto-taxonomy to make it more correct and complete, so to possibly decrease even more the intent recognition error rates.

We have demonstrated in this work that combining exemplar and symbolic ways of defining classes can have a positive impact in the performance of the recognition system. This was done in the context of conversational systems where developers fortuitously embed such alternative descriptions of classes in their name identifiers. We believe it is possible to find in other machine learning development platforms similar patterns of knowledge embedding.

For example, we know that it is common for people to use similar taxonomic structures when naming file and e-mail folders, giving names to functions and variables in programs and data, and writing comments into *Jupyter* notebooks. Also, ML platforms can further foster the use of metadata by developers by explicitly asking them to input, besides exemplars, categorical or textual descriptions of the classes.

As we move along the path of creating such neuro-symbolic systems, not only we should expect that the job of developers becomes easier, as they follow their own cultural and linguistic practices, but also that machines became better in recognizing those classes accurately. Using multiple forms of class definitions can be a winning proposition for both ML systems and their developers.

Ethical Issues

The ChatWorks dataset was composed only of workspaces in which the developers explicitly opted-in to share their code and content for research and development purposes with the company which owns the platform. Those workspaces were shared by the company with the authors of this paper with a clear condition of not publicly sharing their contents and publishing only aggregated results or in an anonymous form. We do not see any specific impact of those limitations in the results of our research but they preclude easy forms of replication of our results with that dataset. To better enable reproducibility, we presented the analysis of the public Larson dataset and shared the intent proto-taxonomy we created manually from its original intent structure in appendix C.

References

Zeynep Akata, Florent Perronnin, Zaid Harchaoui, and Cordelia Schmid. 2016. Label-embedding for image classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(7):1425–1438.

Zeynep Akata, Scott Reed, Daniel Walter, Honglak Lee, and Bernt Schiele. 2015. Evaluation of output embeddings for fine-grained image classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Yoshua Bengio. 2017. The consciousness prior.

Tarek R. Besold, Artur d'Avila Garcez, Sebastian Bader, Howard Bowman, Pedro Domingos, Pascal Hitzler, Kai-Uwe Kuehnberger, Luis C. Lamb, Daniel Lowd, Priscila Machado Vieira Lima, Leo de Penning, Gadi Pinkas, Hoifung Poon, and Gerson Zaverucha. 2017. Neural-symbolic learning and reasoning: A survey and interpretation.

Geoffrey C Bowker and Susan Leigh Star. 2000. *Sorting things out: Classification and its consequences*. MIT press.

Paulo Cavalin, Victor Henrique Alves Ribeiro, Ana Appel, and Claudio Pinhanez. 2020. Improving out-of-scope detection in intent classification by using embeddings of the word graph space of the classes. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 3952–3961, Online. Association for Computational Linguistics.

Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St. John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, Brian Strope, and Ray Kurzweil. 2018. Universal sentence encoder for English. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 169–174, Brussels, Belgium. Association for Computational Linguistics.

Y. Chen, D. Hakkani-Tür, and X. He. 2016. Zero-shot learning of intent embeddings for expansion by convolutional deep structured semantic models. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6045–6049.

Andrea Civan, William Jones, Predrag Klasnja, and Harry Bruce. 2008. Better to organize personal information by folders or by tags?: The devil is in the details. *Proceedings of the American Society for Information Science and Technology*, 45(1):1–13.

Luc De Raedt, R. Manhaeve, S. Dumancic, Thomas Demeester, and A. Kimmig. 2019. Neuro-symbolic = neural + logical + probabilistic. In *Proceedings of the 2019 International Workshop on Neural- Symbolic Learning and Reasoning*, page 4.

Emile Durkheim and Marcel Mauss. 1963. *Primitive Classification*, volume 273. University of Chicago Press.

Lisa Ehrlinger and Wolfram Wöß. 2016. Towards a definition of knowledge graphs. *SEMANTiCS (Posters, Demos, SuCCESS)*, 48:1–4.

Marco Fossati, Dimitris Kontokostas, and Jens Lehmann. 2015. Unsupervised learning of an extensive and usable taxonomy for dbpedia. In *Proceedings of the 11th International Conference on Semantic Systems*, SEM '15. ACM.

Artur d'Avila Garcez, Marco Gori, Luis C. Lamb, Luciano Serafini, Michael Spranger, and Son N. Tran. 2019. Neural-symbolic computing: An effective methodology for principled integration of machine learning and reasoning. *Journal of Applied Logics*, 6(4):611–631.

Drew Hudson and Christopher D Manning. 2019a. Learning by abstraction: The neural state machine. In *Advances in Neural Information Processing Systems*, pages 5903–5916.

Drew Hudson and Christopher D Manning. 2019b. Learning by abstraction: The neural state machine. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 5903–5916. Curran Associates, Inc.

Shaoxiong Ji, Shirui Pan, Erik Cambria, Pekka Marttinen, and Philip S. Yu. 2020. A survey on knowledge graphs: Representation, acquisition and applications.

Daniel Jurafsky and James H. Martin. 2009. *Speech and Language Processing (2nd Edition)*. Prentice-Hall, Inc., USA.

Dimitri Kartsaklis, Mohammad Taher Pilehvar, and Nigel Collier. 2018. Mapping text to knowledge graph entities using multi-sense LSTMs. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 1959–1970, Brussels, Belgium. Association for Computational Linguistics.

Stephen M Kosslyn. 2006. You can play 20 questions with nature and win: Categorical versus coordinate spatial relations as a case study. *Neuropsychologia*, 44(9):1519–1523.

Stefan Larson, Anish Mahendran, Joseph J. Peper, Christopher Clarke, Andrew Lee, Parker Hill, Jonathan K. Kummerfeld, Kevin Leach, Michael A. Laurenzano, Lingjia Tang, and Jason Mars. 2019. An evaluation dataset for intent classification and out-of-scope prediction. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 1311–1316, Hong Kong, China. Association for Computational Linguistics.

Robin Manhaeve, Sebastijan Dumancic, Angelika Kimmig, Thomas Demeester, and Luc De Raedt. 2018. Deepproblog: Neural probabilistic logic programming. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31*, pages 3749–3759. Curran Associates, Inc.

Christopher Manning and Hinrich Schutze. 1999. *Foundations of statistical natural language processing*. MIT press.

Jiayuan Mao, Chuang Gan, Pushmeet Kohli, Joshua B. Tenenbaum, and Jiajun Wu. 2019. The Neuro-Symbolic Concept Learner: Interpreting Scenes, Words, and Sentences From Natural Supervision. In *International Conference on Learning Representations*.

Rodney Needham. 1979. *Symbolic classification*. Goodyear Publishing Company.

Allen Newell. 1973. You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W.G. Chase, editor, *Visual information processing*. Academic Press.

Alessandro Oltramari, Jonathan Francis, Cory Henson, Kaixin Ma, and Ruwan Wickramarachchi. 2020. Neuro-symbolic architectures for context understanding.

Mark Palatucci, Dean Pomerleau, Geoffrey E Hinton, and Tom M Mitchell. 2009. Zero-shot learning with semantic output codes. In *Advances in Neural Information Processing Systems*, volume 22. Curran Associates, Inc.

Emilio Parisotto, Abdel-Rahman Mohamed, Rishabh Singh, Lihong Li, Dengyong Zhou, and Pushmeet Kohli. 2017. Neuro-symbolic program synthesis. In *International Conference on Learning Representations (ICLR)*.

Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. 2014. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 701–710.

Claudio Santos Pinhanez, Heloisa Candello, Paulo Cavalin, Mauro Carlos Pichiliani, Ana Paula Appel, Victor Henrique Alves Ribeiro, Julio Nogima, Maira de Bayser, Melina Guerra, Henrique Ferreira, et al. 2021. Integrating machine learning data with symbolic knowledge from collaboration practices of curators to improve conversational systems. In *Proceedings of the 2021 ACM Conference on Human Factors in Computing Systems (CHI'21)*, pages 1–13.

Whitman Richards. 1982. How to play twenty questions with nature and win.

Richard Socher, Milind Ganjoo, Christopher D Manning, and Andrew Ng. 2013. Zero-shot learning through cross-modal transfer. In *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc.

Ming Tan, Yang Yu, Haoyu Wang, Dakuo Wang, Saloni Potdar, Shiyu Chang, and Mo Yu. 2019. Out-of-domain detection for low-resource text classification tasks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3566–3572, Hong Kong, China. Association for Computational Linguistics.

Joshua B. Tenenbaum, Charles Kemp, Thomas L. Griffiths, and Noah D. Goodman. 2011. How to grow a mind: Statistics, structure, and abstraction. *Science*, 331(6022):1279–1285.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.

Wei Wang, Vincent W Zheng, Han Yu, and Chunyan Miao. 2019. A survey of zero-shot learning: Settings, methods, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(2):1–37.

Steve Whittaker, Tara Matthews, Julian Cerruti, Hernan Badenes, and John Tang. 2011. Am I wasting my time organizing email? a study of email refinding. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, page 3449–3458, New York, NY, USA. Association for Computing Machinery.

Bai Yang, Zhang Liping, and Zhao Fengrong. 2019. A survey on research of code comment. In *Proceedings of the 2019 3rd International Conference on Management Engineering, Software Engineering and Service Sciences*, ICMSS 2019, page 45–51, New York, NY, USA. Association for Computing Machinery.

A An Algorithm to Extract Intent Proto-Taxonomies from NameIds

In our previous work (Pinhanez et al., 2021) we described an algorithm to mine proto-taxonomies from the nameIds of a workspace. For completeness, we include it here. It consists of three steps:

1. Finding the best separator to split the name into a sequence of *concepts*;
2. Splitting the nameIds with the selected separator.
3. Generating an intent proto-taxonomy using the terms split by the best separator as concepts and considering consecutive concepts in a nameId as having a link between them.

In order to find the best separator, our algorithm first calculates the *perplexity* of the bag of concepts using each separator. Perplexity is a measure of uncertainty for a given sequence of words (or concepts) appearing in a language model (Manning and Schutze, 1999). For that, we build probabilistic language models based on *bigrams* using the *maximum likelihood estimator* as in (Jurafsky and Martin, 2009, Chapter 3), and then compute the average perplexity using a standard *leave-one-intent-out* evaluation scheme. The separator which minimizes perplexity is chosen as the separator for the workspace.

B Intent Proto-Taxonomies as a Common Practice in ChatWorks

In (Pinhanez et al., 2021) we evaluated the use of proto-taxonomies by the curators of the ChatWorks dataset, by developing a metric, called *taxonomy rate*, which is the ratio of the number of nameIds which have a proto-taxonomy embedded in it to the total number of intents of the workspace.

To determine whether a given nameId has a proto-taxonomy embedded in it, we considered three criteria: (1) the nameId has two or more concepts; (2) there is at least one other nameId which has at least one identical concept at the exact same level; and (3) the concept does not appear in all nameIds of the workspace in that level. After all the nameIds of a workspace were determined as having an embedded intent or not, the *taxonomy rate* was calculated by considering the ratio between the number of intents with a taxonomic structure and the number of intents of the workspace.

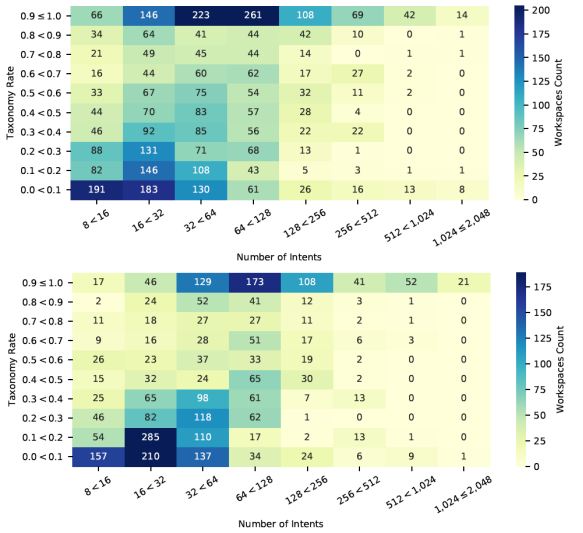

Figure 7: Distribution of the number of workspaces according to the taxonomy rate of the 3,840 English (top) and 2,895 Portuguese (bottom) workspaces. Notice that the X axis is in logarithmic scale.

A taxonomy rate of above 10% was observed in 76% of the English workspaces and in 80% the Portuguese ones. Almost 52% of all workspaces, in English, and 38% in Portuguese, had a taxonomy rate larger than 50%. Moreover, 16% and 20% of the workspaces, for English and Portuguese respectively, had a very high taxonomy rate from 90% to 100%. Considering both the English and Portuguese workspaces, about 70% of them have a taxonomy rate above 20% and more than half of the workspaces above 64 intents have a taxonomy of 50% or more.

Figure 7 shows how the workspaces are distributed considering both the number of intents (X axis) and the taxonomy rate (Y axis). Notice that the distribution, in both languages, follows a sort of a "step" function where, as the threshold between 32 and 64 intents is crossed, the majority of the workspaces has more than 50% of taxonomy rate. More details can be found in (Pinhanez et al., 2021).

C The Intent Proto-Taxonomy Created for the Larson Dataset

In tables 1 and 2, we list the taxonomy which was manually created for the Larson dataset, with the original nameId on the left and the created taxonomic representation on the right, represented as a string of concepts separated by spaces.

nameId	Concepts
accept_reservations	accept reservation
account_blocked	account blocked
alarm	set alarm
application_status	application status
apr	what month
are_you_a_bot	you bot
balance	what balance
bill_balance	what bill balance
bill_due	when bill due
book_flight	book flight
book_hotel	book hotel
calculator	calculate
calendar	calendar
calendar_update	calendar update
calories	calories dish
cancel	cancel action
cancel_reservation	cancel reservation
car_rental	car rental
card_declined	card declined
carry_on	carry-on rule
change_accent	change accent
change_ai_name	change bot name
change_language	change language
change_speed	change speed
change_user_name	change user name
change_volume	change volume
confirm_reservation	confirm reservation
cook_time	cook time
credit_limit	credit limit
credit_limit_change	credit limit change
credit_score	credit score
current_location	what current location
damaged_card	damaged card
date	what date
definition	definition
direct_deposit	direct deposit
directions	what direction
distance	what distance
do_you_have_pets	do you have pet
exchange_rate	exchange rate
expiration_date	expiration date
find_phone	find phone
flight_status	flight status
flip_coin	flip coin
food_last	food last
freeze_account	block account
fun_fact	fun fact
gas	gas level
gas_type	gas type
goodbye	goodbye
greeting	greeting
how_busy	how busy
how_old_are_you	how old you
improve_credit_score	improve credit score
income	what income
ingredient_substitution	ingredient substitution
ingredients_list	ingredient list
insurance	insurance benefit
insurance_change	insurance change
interest_rate	interest rate
international_fees	international fee
international_visa	international visa
jump_start	jump start
last_maintenance	last maintenance
lost_luggage	lost luggage
make_call	make call
maybe	maybe
meal_suggestion	meal suggestion
meaning_of_life	what meaning life
measurement_conversion	measurement conversion
meeting_schedule	meeting schedule
min_payment	minimum payment
mpg	what mpg
new_card	apply card
next_holiday	next holiday
next_song	next song
no	no
nutrition_info	nutrition info
oil_change_how	how change oil
oil_change_when	when change oil
order	order shopping
order_checks	order check
order_status	order status
pay_bill	pay bill
payday	when payday
pin_change	change pin
play_music	play music
plug_type	what plug type
pto_balance	pto balance
pto_request	pto request
	. . .

Table 1: The taxonomy created for the Larson dataset.

nameId	Concepts
	. . .
pto_request_status	pto request status
pto_used	pto used
recipe	recipe dish
redeem_rewards	redeem reward
reminder	reminder action
reminder_update	reminder update
repeat	repeat action
replacement_card_duration	replacement card duration
report_fraud	report fraud
report_lost_card	report lost card
reset_settings	reset setting
restaurant_reservation	restaurant reservation
restaurant_reviews	restaurant review
restaurant_suggestion	restaurant suggestion
rewards_balance	reward balance
roll_dice	roll dice
rollover_401k	rollover 401k
routing	find routing
schedule_maintenance	schedule maintenance
schedule_meeting	schedule meeting
share_location	share location
shopping_list	shopping list
shopping_list_update	shopping list update
smart_home	smart home
spelling	spelling word
spending_history	spending history
sync_device	sync device
taxes	what taxes
tell_joke	tell joke
text	text person
thank_you	thank
time	what time
timer	set timer
timezone	set timezone
tire_change	tire change
tire_pressure	tire pressure
todo_list	todo list
todo_list_update	todo list update
traffic	what traffic
transactions	card transaction
transfer	transfer account
translate	translate word
travel_alert	travel alert
travel_notification	travel notification
travel_suggestion	travel suggestion
uber	get uber
update_playlist	update playlist
user_name	user name
vaccines	what vaccine
w2	get w2
weather	what weather
what_are_your_hobbies	what hobby
what_can_i_ask_you	what ask you
what_is_your_name	what name
what_song	what song
where_are_you_from	where you from
whisper_mode	whisper mode
who_do_you_work_for	who you work
who_made_you	who made you
yes	yes
	. . .

Table 2: The taxonomy created for the Larson dataset (cont.).

D Results of the Proposed Algorithms in the ChatWorks Dataset

Table 3 shows the results of the proposed three algorithms and the baseline for each workspace in the ChatWorks dataset. For each metrics, EER, FAR, and ISER, the table shows the average accuracy and the standard deviation over the experiments made with the 20 random splits. The table also lists information characterizing the workspaces: number of intents, number of exemplars, the average number of exemplars per intent, taxonomy rate, the average depth of the taxonomy graph, and the number of different concepts.

WORKSPACE ID	USE - EER	USE - EER(std)	USE - FAR	USE - FAR(std)	USE - ISER	USE - ISER(std)	USE+T - EER	USE+T - EER(std)	USE+T - FAR	USE+T - FAR(std)	USE+T - ISER	USE+T - ISER(std)	USE+S - EER	USE+S - EER(std)	USE+S - FAR	USE+S - FAR(std)	USE+S - ISER	USE+S - ISER(std)	USE+C - EER	USE+C - EER(std)	USE+C - FAR	USE+C - FAR(std)	USE+C - ISER	USE+C - ISER(std)	intents	exemplars	examples_avg	taxonomy_rate	tax_depth_avg	concepts
0	0,268	0,006	0,390	0,090	0,300	0,006	0,267	0,012	0,286	0,079	0,337	0,006	0,274	0,010	0,286	0,091	0,418	0,013	0,250	0,009	0,311	0,088	0,329	0,010	110	5655	51,409	0,709	2,545	161
1	0,224	0,011	0,272	0,151	0,141	0,010	0,285	0,014	0,232	0,149	0,195	0,013	0,222	0,013	0,367	0,127	0,202	0,014	0,272	0,014	0,203	0,115	0,168	0,010	190	3220	16,947	0,942	4,768	297
2	0,598	0,095	0,553	0,476	0,057	0,012	0,862	0,145	0,450	0,497	0,102	0,043	0,876	0,134	0,450	0,497	0,078	0,016	0,890	0,139	0,450	0,497	0,065	0,019	17	506	29,765	0,529	3,000	30
3	0,149	0,018	0,355	0,287	0,158	0,014	0,238	0,031	0,172	0,163	0,240	0,041	0,178	0,022	0,321	0,267	0,192	0,025	0,163	0,018	0,337	0,255	0,196	0,019	29	534	18,414	0,414	1,690	43
4	0,098	0,039	0,362	0,306	0,046	0,011	0,128	0,082	0,320	0,334	0,102	0,027	0,110	0,081	0,357	0,319	0,073	0,008	0,104	0,020	0,265	0,327	0,065	0,011	41	1463	35,683	0,463	2,146	56
5	0,140	0,076	0,402	0,261	0,084	0,006	0,175	0,064	0,568	0,206	0,109	0,007	0,174	0,063	0,357	0,216	0,091	0,007	0,140	0,080	0,247	0,295	0,084	0,007	22	7380	335,455	0,636	2,273	24
6	0,128	0,018	0,301	0,272	0,091	0,010	0,296	0,042	0,481	0,311	0,306	0,039	0,129	0,032	0,410	0,330	0,142	0,025	0,161	0,024	0,111	0,216	0,116	0,015	23	422	18,348	0,957	2,652	33
7	0,329	0,015	0,408	0,124	0,487	0,012	0,366	0,019	0,501	0,122	0,524	0,023	0,328	0,021	0,396	0,169	0,508	0,017	0,318	0,019	0,368	0,120	0,475	0,015	122	1921	15,746	0,385	2,238	170
8	0,125	0,071	0,462	0,362	0,079	0,008	0,112	0,019	0,507	0,363	0,077	0,011	0,123	0,019	0,514	0,350	0,074	0,009	0,138	0,033	0,518	0,367	0,137	0,063	26	5886	226,385	0,346	1,577	32
9	0,219	0,008	0,328	0,087	0,249	0,010	0,269	0,010	0,397	0,079	0,345	0,010	0,259	0,010	0,255	0,072	0,344	0,012	0,279	0,011	0,253	0,079	0,372	0,010	322	4496	13,963	0,963	3,283	331
10	0,601	0,094	0,500	0,500	0,160	0,034	0,639	0,091	0,500	0,500	0,333	0,042	0,743	0,108	0,500	0,500	0,222	0,058	0,149	0,066	0,660	0,441	0,195	0,035	15	339	22,600	0,867	2,000	18
11	0,176	0,013	0,329	0,099	0,130	0,010	0,167	0,018	0,362	0,189	0,164	0,013	0,180	0,032	0,362	0,237	0,180	0,012	0,154	0,019	0,279	0,132	0,132	0,009	47	2721	57,894	0,681	1,957	39
12	0,256	0,019	0,449	0,281	0,215	0,021	0,510	0,043	0,769	0,272	0,530	0,040	0,686	0,041	0,050	0,218	0,280	0,038	0,273	0,028	0,593	0,341	0,295	0,025	31	462	14,903	0,935	2,774	36
13	0,197	0,023	0,656	0,275	0,136	0,019	0,577	0,044	0,350	0,477	0,394	0,038	0,198	0,032	0,627	0,325	0,234	0,030	0,265	0,024	0,465	0,407	0,165	0,014	13	707	54,385	0,846	2,615	20
14	0,174	0,008	0,192	0,067	0,103	0,006	0,172	0,013	0,293	0,086	0,160	0,006	0,175	0,011	0,300	0,106	0,188	0,009	0,146	0,009	0,301	0,113	0,124	0,008	290	3255	11,224	0,979	2,431	344
15	0,015	0,011	0,211	0,362	0,019	0,003	0,125	0,055	0,863	0,186	0,089	0,025	0,072	0,045	0,832	0,228	0,038	0,009	0,398	0,027	0,844	0,246	0,372	0,044	42	627	14,929	0,857	2,738	52
16	0,298	0,039	0,400	0,490	0,112	0,024	0,288	0,048	0,670	0,457	0,221	0,046	0,280	0,055	0,460	0,490	0,156	0,030	0,236	0,044	0,400	0,490	0,124	0,030	44	597	13,568	0,341	2,364	68
17	0,104	0,021	0,279	0,119	0,088	0,009	0,121	0,017	0,204	0,111	0,109	0,011	0,113	0,020	0,221	0,135	0,114	0,012	0,103	0,009	0,202	0,153	0,133	0,013	143	5240	36,643	0,636	1,986	158
18	0,739	0,072	0,563	0,487	0,107	0,012	0,758	0,069	0,550	0,497	0,201	0,049	0,767	0,056	0,550	0,497	0,192	0,030	0,287	0,045	0,702	0,358	0,130	0,021	10	360	36,000	0,300	1,900	10
19	0,066	0,012	0,303	0,276	0,065	0,031	0,105	0,014	0,217	0,207	0,112	0,046	0,119	0,017	0,278	0,265	0,086	0,034	0,143	0,016	0,171	0,170	0,070	0,041	57	925	16,228	0,965	3,088	79
20	0,199	0,013	0,330	0,119	0,145	0,010	0,194	0,014	0,353	0,134	0,210	0,013	0,251	0,021	0,981	0,033	0,217	0,017	0,216	0,020	0,185	0,109	0,194	0,011	91	2969	32,626	0,692	2,429	114
21	0,170	0,008	0,314	0,099	0,150	0,006	0,215	0,010	0,169	0,083	0,167	0,011	0,147	0,012	0,341	0,094	0,200	0,010	0,172	0,008	0,261	0,100	0,144	0,006	76	4620	60,789	0,882	3,092	119
22	0,221	0,019	0,327	0,149	0,215	0,011	0,223	0,019	0,299	0,146	0,240	0,010	0,216	0,019	0,314	0,130	0,265	0,017	0,179	0,017	0,360	0,147	0,258	0,015	107	1583	14,794	0,430	1,794	141
23	0,073	0,021	0,223	0,336	0,050	0,008	0,159	0,026	0,206	0,302	0,045	0,015	0,149	0,043	0,183	0,315	0,068	0,022	0,163	0,040	0,148	0,260	0,069	0,014	49	546	11,143	0,429	3,837	125
24	0,211	0,031	0,225	0,402	0,107	0,021	0,456	0,063	0,475	0,460	0,343	0,058	0,430	0,061	0,300	0,458	0,166	0,045	0,267	0,043	0,350	0,450	0,090	0,028	32	376	11,750	0,438	2,469	48
25	0,167	0,006	0,278	0,083	0,131	0,004	0,209	0,006	0,171	0,075	0,156	0,005	0,138	0,011	0,606	0,092	0,162	0,007	0,361	0,435	0,375	0,414	0,996	0,001	358	8224	22,972	0,902	4,648	495
26	0,116	0,017	0,370	0,307	0,101	0,011	0,210	0,032	0,651	0,287	0,182	0,024	0,155	0,043	0,767	0,258	0,125	0,017	0,210	0,028	0,249	0,308	0,105	0,013	25	655	26,200	0,920	2,720	42
27	0,208	0,010	0,358	0,212	0,216	0,012	0,231	0,026	0,399	0,270	0,248	0,029	0,221	0,030	0,426	0,283	0,249	0,019	0,198	0,028	0,307	0,246	0,223	0,023	101	1272	12,594	0,604	2,752	158
28	0,140	0,073	0,534	0,378	0,073	0,006	0,113	0,025	0,530	0,394	0,075	0,007	0,126	0,019	0,535	0,374	0,071	0,006	0,149	0,034	0,581	0,355	0,143	0,061	26	5890	226,538	0,346	1,577	32
29	0,374	0,034	0,567	0,458	0,156	0,030	0,526	0,078	0,500	0,500	0,446	0,082	0,236	0,085	0,725	0,402	0,247	0,073	0,472	0,052	0,525	0,487	0,168	0,033	21	349	16,619	0,476	2,810	38
30	0,245	0,010	0,370	0,080	0,345	0,006	0,249	0,009	0,366	0,107	0,376	0,009	0,259	0,013	0,355	0,131	0,444	0,015	0,240	0,010	0,328	0,120	0,375	0,010	153	3691	24,124	0,536	2,359	202
31	0,213	0,020	0,377	0,266	0,243	0,017	0,255	0,022	0,227	0,204	0,265	0,018	0,220	0,032	0,386	0,260	0,271	0,023	0,294	0,028	0,299	0,182	0,329	0,026	59	745	12,627	0,339	2,458	117
32	0,222	0,012	0,374	0,202	0,316	0,012	0,307	0,024	0,425	0,195	0,486	0,028	0,244	0,030	0,168	0,121	0,456	0,021	0,268	0,021	0,401	0,150	0,371	0,020	97	1111	11,454	0,773	2,804	132
33	0,063	0,026	0,419	0,394	0,075	0,017	0,246	0,059	0,509	0,416	0,204	0,050	0,190	0,042	0,479	0,425	0,155	0,045	0,205	0,050	0,244	0,346	0,088	0,020	23	394	17,130	0,478	2,348	42
34	0,159	0,032	0,498	0,338	0,220	0,025	0,281	0,071	0,325	0,400	0,333	0,050	0,239	0,062	0,410	0,402	0,276	0,039	0,271	0,054	0,408	0,439	0,230	0,043	16	179	11,188	0,625	4,563	48
35	0,298	0,162	0,336	0,174	0,333	0,153	0,273	0,014	0,307	0,128	0,362	0,017	0,268	0,018	0,419	0,142	0,377	0,019	0,230	0,013	0,346	0,122	0,310	0,009	81	2475	30,556	0,568	1,914	99
36	0,220	0,016	0,326	0,261	0,223	0,014	0,231	0,031	0,418	0,267	0,238	0,025	0,252	0,027	0,383	0,291	0,236	0,018	0,252	0,027	0,347	0,252	0,238	0,016	60	755	12,583	0,350	2,450	118
37	0,196	0,016	0,236	0,107	0,182	0,014	0,347	0,026	0,449	0,166	0,490	0,028	0,310	0,019	0,991	0,029	0,278	0,022	0,416	0,022	0,845	0,144	0,424	0,023	82	858	10,463	0,976	2,512	90
38	0,546	0,077	0,517	0,462	0,196	0,046	0,408	0,085	0,300	0,458	0,580	0,098	0,369	0,074	0,500	0,447	0,407	0,088	0,492	0,055	0,475	0,460	0,223	0,041	23	252	10,957	0,565	2,957	37
39	0,176	0,012	0,192	0,190	0,141	0,013	0,208	0,034	0,213	0,204	0,180	0,027	0,184	0,028	0,258	0,204	0,199	0,023	0,156	0,019	0,281	0,216	0,154	0,014	39	680	17,436	0,385	2,256	66
40	0,147	0,025	0,157	0,261	0,122	0,019	0,246	0,045	0,268	0,267	0,186	0,035	0,261	0,036	0,189	0,280	0,177	0,041	0,237	0,035	0,285	0,346	0,176	0,039	64	734	11,469	0,344	2,406	107
41	0,868	0,073	0,750	0,433	0,085	0,029	0,623	0,136	0,750	0,433	0,350	0,123	0,809	0,076	0,750	0,433	0,154	0,075	0,468	0,092	0,775	0,402	0,085	0,029	17	292	17,176	0,176	3,235	32
42	0,192	0,010	0,240	0,114	0,157	0,008	0,152	0,011	0,269	0,135	0,167	0,010	0,178	0,021	0,385	0,155	0,192	0,011	0,169	0,011	0,382	0,124	0,489	0,022	100	3222	32,220	0,650	2,590	136
43	0,235	0,008	0,277	0,098	0,230	0,010	0,253	0,013	0,364	0,086	0,300	0,014	0,259	0,017	0,331	0,144	0,319	0,015	0,284	0,013	0,277	0,095	0,355	0,009	211	2957	14,014	0,967	2,934	218
44	0,176	0,017	0,261	0,167	0,102	0,008	0,168	0,015	0,293	0,163	0,151	0,018	0,234	0,025	0,188	0,156	0,160	0,020	0,194	0,024	0,261	0,158	0,117	0,013	52	984	18,923	0,904	2,596	91
45	0,115	0,020	0,254	0,116	0,122	0,011	0,140	0,009	0,183	0,155	0,131	0,013	0,132	0,010	0,200	0,172	0,137	0,012	0,134	0,009	0,215	0,132	0,160	0,015	174	5616	32,276	0,621	1,891	196
46	0,223	0,021	0,240	0,210	0,125	0,013	0,201	0,030	0,351	0,256	0,165	0,020	0,271	0,026	0,313	0,236	0,195	0,021	0,204	0,020	0,281	0,210	0,147	0,018	53	538	10,151	0,811	1,717	56
47	0,243	0,011	0,315	0,092	0,293	0,010	0,213	0,009	0,301	0,096	0,327	0,013	0,211	0,016	0,357	0,122	0,365	0,012	0,214	0,011	0,338	0,105	0,319	0,011	292	3010	10,308	0,716	8,627	719
48	0,238	0,020	0,367	0,266	0,180	0,009	0,246	0,029	0,322	0,228	0,229	0,018	0,263	0,025	0,321	0,247	0,250	0,025	0,305	0,033	0,304	0,209	0,275	0,028	60	766	12,767	0,350	2,450	118
49	0,150	0,011	0,218	0,090	0,139	0,016	0,164	0,020	0,313	0,137	0,192	0,014	0,148	0,019	0,311	0,147	0,213	0,013	0,207	0,017	0,149	0,095	0,184	0,017	87	1195	13,736	1,000	3,805	124
50	0,164	0,010	0,443	0,156	0,128	0,007	0,187	0,013	0,350	0,181	0,145	0,012	0,167	0,014	0,385	0,223	0,146	0,014	0,196	0,010	0,232	0,127	0,125	0,011	33	3035	91,970	0,455	2,939	57
51	0,833	0,068	0,550	0,497	0,020	0,007	0,782	0,063	0,550	0,497	0,183	0,052	0,852	0,069	0,550	0,497	0,101	0,048	0,210	0,049	0,617	0,475	0,027	0,013	14	287	20,500	0,643	3,214	31
52	0,240	0,024	0,299	0,292	0,185	0,014	0,208	0,037	0,347	0,266	0,229	0,028	0,233	0,031	0,255	0,270	0,250	0,022	0,229	0,030	0,297	0,319	0,233	0,016	47	506	10,766	0,319	2,340	90
53	0,093	0,006	0,168	0,070	0,056	0,004	0,124	0,009	0,093	0,065	0,068	0,006	0,102	0,008	0,131	0,076	0,064	0,004	0,086	0,006	0,167	0,070	0,081	0,006	273	5012	18,359	0,546	2,875	467
54	0,190	0,018	0,292	0,237	0,158	0,015	0,360	0,059	0,439	0,331	0,454	0,048	0,242	0,027	0,345	0,311	0,224	0,021	0,181	0,026	0,335	0,258	0,163	0,020	26	804	30,923	0,962	1,962	30
55	0,233	0,011	0,292	0,106	0,221	0,009	0,237	0,013	0,377	0,117	0,285	0,010	0,241	0,012	0,368	0,092	0,308	0,011	0,252	0,009	0,258	0,083	0,345	0,010	192	2713	14,130	0,958	2,917	195
56	0,636	0,033	0,500	0,500	0,092	0,019	0,685	0,090	0,500	0,500	0,278	0,087	0,175	0,049	0,604	0,445	0,125	0,030	0,173	0,050	0,613	0,441	0,121	0,021	28	658	23,500	0,143	2,893	51
57	0,341	0,035	0,325	0,455	0,247	0,026	0,461	0,049	0,500	0,450	0,510	0,067	0,450	0,059	0,292	0,428	0,333	0,061	0,318	0,052	0,458	0,474	0,247	0,029	33	409	12,394	0,545	2,455	48
58	0,149	0,044	0,490	0,284	0,099	0,006	0,146	0,028	0,397	0,269	0,107	0,009	0,153	0,022	0,423	0,259	0,104	0,005	0,160	0,020	0,341	0,300	0,100	0,005	39	5843	149,821	0,821	3,308	66
59	0,230	0,016	0,246	0,107	0,131	0,008	0,215	0,023	0,325	0,145	0,180	0,016	0,256	0,017	0,294	0,109	0,265	0,014	0,188	0,018	0,339	0,148	0,174	0,012	99	1386	14,000	0,697	2,747	174
60	0,156	0,013	0,258	0,098	0,154	0,007	0,166	0,010	0,259	0,120	0,225	0,010	0,154	0,014	0,209	0,088	0,231	0,006	0,151	0,011	0,248	0,099	0,177	0,006	164	2870	17,500	1,000	3,262	208
61	0,160	0,011	0,309	0,134	0,087	0,007	0,177	0,012	0,205	0,106	0,126	0,011	0,152	0,010	0,275	0,134	0,121	0,010	0,200	0,012	0,170	0,130	0,116	0,009	103	2257	21,913	0,767	3,068	149
62	0,092	0,014	0,180	0,230	0,144	0,014	0,427	0,031	0,354	0,265	0,483	0,045	0,246	0,046	0,175	0,268	0,216	0,030	0,141	0,031	0,293	0,246	0,154	0,016	27	332	12,296	0,815	2,074	32
63	0,208	0,011	0,342	0,125	0,137	0,009	0,205	0,018	0,355	0,170	0,203	0,016	0,269	0,020	0,980	0,048	0,234	0,018	0,245	0,014	0,197	0,116	0,209	0,014	94	2849	30,309	0,702	2,426	117
64	0,024	0,007	0,243	0,238	0,014	0,004	0,036	0,006	0,205	0,206	0,030	0,006	0,036	0,012	0,172	0,183	0,024	0,008	0,037	0,006	0,178	0,199	0,022	0,005	42	3238	77,095	0,429	2,357	68
65	0,492	0,074	0,850	0,357	0,295	0,075	0,596	0,084	0,700	0,458	0,366	0,120	0,642	0,084	0,700	0,458	0,319	0,106	0,654	0,076	0,700	0,458	0,308	0,088	16	293	18,313	0,250	2,875	30
66	0,171	0,010	0,287	0,121	0,124	0,006	0,216	0,014	0,163	0,105	0,148	0,008	0,159	0,011	0,458	0,132	0,181	0,007	0,146	0,009	0,264	0,123	0,140	0,007	77	3075	39,935	0,584	3,286	111
67	0,273	0,022	0,286	0,234	0,196	0,014	0,264	0,023	0,346	0,213	0,242	0,020	0,258	0,021	0,290	0,179	0,263	0,025	0,264	0,026	0,305	0,227	0,210	0,016	66	750	11,364	0,424	2,167	103
68	0,070	0,023	0,550	0,404	0,077	0,008	0,234	0,055	0,666	0,366	0,206	0,043	0,254	0,059	0,437	0,468	0,125	0,030	0,218	0,055	0,443	0,466	0,099	0,016	12	256	21,333	0,417	2,583	27
69	0,215	0,009	0,348	0,111	0,206	0,007	0,267	0,024	0,474	0,107	0,379	0,033	0,216	0,011	0,379	0,130	0,258	0,014	0,198	0,012	0,329	0,102	0,203	0,008	32	4963	155,094	0,531	2,000	30
70	0,222	0,024	0,636	0,326	0,172	0,012	0,289	0,023	0,571	0,315	0,269	0,027	0,260	0,012	0,480	0,373	0,167	0,014	0,257	0,013	0,469	0,366	0,161	0,010	15	5986	399,067	0,133	0,000	31
71	0,185	0,011	0,255	0,138	0,104	0,005	0,194	0,016	0,217	0,091	0,140	0,013	0,183	0,018	0,356	0,157	0,157	0,014	0,216	0,015	0,129	0,087	0,121	0,011	98	1815	18,520	0,918	2,755	120
72	0,248	0,015	0,321	0,116	0,232	0,010	0,224	0,014	0,344	0,110	0,300	0,012	0,353	0,019	0,957	0,062	0,321	0,014	0,202	0,016	0,439	0,122	0,300	0,015	102	3076	30,157	0,716	2,431	124
73	0,228	0,028	0,250	0,403	0,090	0,014	0,680	0,059	0,150	0,357	0,288	0,056	0,198	0,060	0,608	0,463	0,202	0,033	0,314	0,044	0,250	0,403	0,160	0,019	22	265	12,045	0,682	2,500	30
74	0,290	0,012	0,436	0,094	0,381	0,011	0,289	0,009	0,293	0,087	0,406	0,012	0,263	0,015	0,302	0,114	0,434	0,010	0,267	0,006	0,327	0,093	0,388	0,012	107	4976	46,505	0,393	4,439	266
75	0,028	0,012	0,215	0,233	0,023	0,007	0,035	0,006	0,195	0,253	0,029	0,005	0,034	0,006	0,176	0,214	0,030	0,009	0,036	0,011	0,140	0,198	0,027	0,006	40	3207	80,175	0,425	2,275	66
76	0,332	0,020	0,476	0,265	0,293	0,018	0,282	0,027	0,421	0,273	0,339	0,022	0,281	0,020	0,319	0,271	0,368	0,023	0,289	0,021	0,364	0,261	0,299	0,022	94	1243	13,223	0,606	2,543	123
77	0,210	0,010	0,311	0,081	0,196	0,007	0,209	0,013	0,349	0,113	0,260	0,011	0,204	0,011	0,445	0,115	0,277	0,013	0,197	0,011	0,315	0,088	0,229	0,008	268	3967	14,802	0,440	2,481	339
78	0,218	0,019	0,356	0,146	0,219	0,010	0,211	0,018	0,320	0,130	0,240	0,013	0,215	0,024	0,225	0,131	0,282	0,012	0,205	0,023	0,277	0,120	0,225	0,014	159	2193	13,792	0,799	2,887	217
79	0,140	0,008	0,233	0,063	0,053	0,005	0,198	0,010	0,112	0,048	0,115	0,012	0,155	0,008	0,162	0,076	0,087	0,011	0,086	0,009	0,203	0,085	0,062	0,005	186	4236	22,774	0,839	2,081	232
80	0,121	0,020	0,290	0,255	0,115	0,013	0,300	0,049	0,399	0,300	0,355	0,037	0,206	0,043	0,861	0,243	0,182	0,034	0,149	0,022	0,248	0,195	0,152	0,024	23	414	18,000	0,957	2,783	44
81	0,950	0,093	0,750	0,433	0,000	0,000	0,850	0,104	0,750	0,433	0,107	0,099	0,942	0,092	0,750	0,433	0,008	0,025	0,950	0,093	0,750	0,433	0,000	0,000	43	806	18,744	0,000	3,372	73
82	0,167	0,011	0,357	0,133	0,156	0,008	0,170	0,011	0,303	0,127	0,184	0,009	0,183	0,030	0,631	0,145	0,217	0,015	0,197	0,012	0,300	0,126	0,202	0,008	84	3713	44,202	0,857	3,667	145
83	0,736	0,053	0,700	0,458	0,110	0,012	0,738	0,122	0,700	0,458	0,231	0,072	0,807	0,083	0,700	0,458	0,153	0,035	0,632	0,083	0,700	0,458	0,335	0,063	10	238	23,800	0,400	2,400	16
84	0,285	0,008	0,415	0,076	0,461	0,006	0,304	0,012	0,392	0,060	0,505	0,010	0,298	0,011	0,407	0,056	0,518	0,011	0,283	0,010	0,400	0,075	0,471	0,008	118	7896	66,915	0,542	1,678	128
85	0,161	0,024	0,222	0,295	0,181	0,018	0,189	0,038	0,208	0,279	0,226	0,026	0,177	0,031	0,312	0,312	0,218	0,036	0,188	0,026	0,272	0,277	0,255	0,028	24	303	12,625	0,292	2,292	50
86	0,212	0,013	0,314	0,121	0,256	0,012	0,215	0,018	0,281	0,083	0,275	0,015	0,270	0,017	0,471	0,161	0,317	0,014	0,244	0,014	0,358	0,112	0,306	0,013	134	2072	15,463	0,567	2,560	160
87	0,939	0,088	0,700	0,458	0,012	0,028	0,796	0,102	0,700	0,458	0,164	0,097	0,479	0,088	0,733	0,429	0,027	0,036	0,479	0,072	0,717	0,438	0,030	0,038	16	303	18,938	0,125	2,375	26
88	0,218	0,023	0,370	0,262	0,234	0,046	0,311	0,027	0,251	0,196	0,234	0,025	0,306	0,023	0,301	0,214	0,257	0,045	0,311	0,036	0,240	0,178	0,262	0,039	58	1082	18,655	1,000	2,672	117
89	0,419	0,053	0,650	0,477	0,055	0,036	0,697	0,125	0,650	0,477	0,276	0,097	0,284	0,078	0,650	0,477	0,125	0,060	0,384	0,050	0,650	0,477	0,091	0,045	30	568	18,933	0,000	3,333	63
90	0,302	0,051	0,664	0,429	0,182	0,038	0,750	0,075	0,550	0,497	0,211	0,034	0,283	0,045	0,605	0,452	0,186	0,028	0,286	0,036	0,684	0,417	0,180	0,026	34	554	16,294	0,088	2,882	52
91	0,291	0,009	0,388	0,083	0,315	0,012	0,266	0,014	0,276	0,091	0,347	0,012	0,279	0,016	0,280	0,083	0,426	0,011	0,251	0,012	0,296	0,076	0,348	0,011	99	3313	33,465	0,677	2,505	150
92	0,193	0,012	0,300	0,208	0,132	0,012	0,226	0,024	0,333	0,208	0,176	0,018	0,237	0,034	0,952	0,114	0,198	0,017	0,206	0,026	0,272	0,134	0,189	0,010	80	1029	12,863	0,863	3,038	115
93	0,583	0,137	0,650	0,477	0,136	0,052	0,617	0,067	0,600	0,490	0,369	0,064	0,375	0,100	0,675	0,455	0,165	0,086	0,425	0,087	0,650	0,477	0,092	0,034	22	306	13,909	0,318	2,636	38
94	0,154	0,008	0,246	0,064	0,130	0,004	0,174	0,012	0,254	0,103	0,172	0,009	0,154	0,013	0,273	0,126	0,167	0,008	0,168	0,009	0,182	0,073	0,146	0,007	221	2857	12,928	0,502	2,303	314
95	0,168	0,017	0,353	0,338	0,123	0,028	0,273	0,038	0,280	0,371	0,230	0,029	0,278	0,031	0,236	0,320	0,211	0,027	0,266	0,035	0,213	0,315	0,196	0,030	26	396	15,231	0,308	2,308	50
96	0,146	0,022	0,257	0,181	0,160	0,016	0,202	0,026	0,270	0,296	0,247	0,031	0,173	0,037	0,135	0,256	0,264	0,036	0,164	0,029	0,265	0,278	0,179	0,021	42	445	10,595	1,000	2,929	67
97	0,167	0,009	0,283	0,114	0,148	0,008	0,188	0,012	0,254	0,103	0,175	0,007	0,172	0,011	0,286	0,102	0,171	0,006	0,174	0,010	0,199	0,099	0,150	0,006	212	2801	13,212	0,472	2,231	306
98	0,245	0,029	0,275	0,402	0,065	0,014	0,453	0,064	0,925	0,238	0,373	0,066	0,345	0,058	0,467	0,461	0,182	0,044	0,231	0,033	0,275	0,402	0,034	0,008	25	321	12,840	0,440	1,760	33
99	0,155	0,019	0,276	0,272	0,131	0,011	0,410	0,054	0,560	0,250	0,514	0,057	0,191	0,032	0,475	0,272	0,246	0,026	0,020	0,005	0,115	0,153	0,968	0,007	25	632	25,280	1,000	2,480	33
100	0,072	0,017	0,224	0,215	0,024	0,007	0,193	0,046	0,143	0,170	0,050	0,018	0,185	0,031	0,127	0,157	0,071	0,028	0,109	0,036	0,230	0,275	0,062	0,013	52	561	10,788	0,404	3,750	132
101	0,244	0,022	0,256	0,190	0,188	0,012	0,220	0,026	0,359	0,236	0,237	0,022	0,206	0,016	0,349	0,198	0,244	0,018	0,235	0,019	0,279	0,168	0,209	0,012	57	905	15,877	0,509	2,421	94
102	0,310	0,027	0,605	0,375	0,467	0,030	0,403	0,037	0,250	0,433	0,576	0,035	0,521	0,023	0,802	0,316	0,521	0,035	0,337	0,027	0,581	0,372	0,484	0,031	15	967	64,467	0,800	1,800	18
103	0,134	0,012	0,240	0,206	0,090	0,007	0,200	0,015	0,201	0,161	0,122	0,010	0,155	0,025	0,182	0,181	0,143	0,019	0,214	0,023	0,188	0,168	0,153	0,021	99	1074	10,848	0,707	2,394	134
104	0,182	0,031	0,602	0,371	0,147	0,027	0,635	0,099	0,350	0,477	0,337	0,075	0,260	0,062	0,548	0,421	0,205	0,034	0,221	0,034	0,446	0,435	0,185	0,038	32	683	21,344	0,125	2,594	50
105	0,202	0,008	0,316	0,066	0,200	0,005	0,217	0,006	0,237	0,053	0,249	0,010	0,252	0,012	0,533	0,160	0,301	0,010	0,257	0,012	0,423	0,143	0,274	0,009	298	5183	17,393	0,946	3,416	381
106	0,088	0,032	0,492	0,437	0,080	0,013	0,133	0,029	0,571	0,445	0,090	0,014	0,129	0,026	0,444	0,457	0,075	0,010	0,119	0,023	0,452	0,459	0,075	0,009	27	1206	44,667	0,259	2,630	49
107	0,169	0,010	0,292	0,086	0,145	0,006	0,246	0,014	0,280	0,108	0,261	0,011	0,258	0,016	0,887	0,056	0,243	0,009	0,329	0,016	0,497	0,093	0,356	0,009	452	5802	12,836	0,998	3,308	523
108	0,278	0,012	0,424	0,130	0,325	0,011	0,291	0,017	0,352	0,111	0,409	0,039	0,265	0,014	0,319	0,138	0,396	0,024	0,277	0,016	0,394	0,145	0,348	0,016	112	4385	39,152	0,411	2,348	197
109	0,223	0,033	0,430	0,361	0,145	0,019	0,360	0,066	0,707	0,362	0,348	0,071	0,250	0,058	1,000	0,000	0,212	0,041	0,185	0,038	0,331	0,367	0,130	0,025	23	284	12,348	0,478	2,000	36
110	0,255	0,012	0,352	0,112	0,243	0,009	0,243	0,015	0,373	0,118	0,275	0,010	0,257	0,016	0,267	0,124	0,305	0,014	0,244	0,014	0,379	0,087	0,262	0,010	174	2826	16,241	0,805	2,868	225
111	0,290	0,033	0,401	0,263	0,443	0,018	0,275	0,032	0,412	0,250	0,475	0,031	0,368	0,047	0,530	0,263	0,514	0,035	0,282	0,031	0,320	0,232	0,448	0,028	23	562	24,435	0,391	3,522	58
112	0,170	0,013	0,307	0,110	0,142	0,009	0,180	0,013	0,332	0,178	0,165	0,013	0,171	0,013	0,217	0,102	0,207	0,015	0,180	0,016	0,277	0,138	0,149	0,010	57	1813	31,807	0,684	3,123	106
113	0,196	0,007	0,305	0,117	0,187	0,009	0,182	0,014	0,235	0,144	0,212	0,010	0,179	0,014	0,241	0,105	0,216	0,010	0,187	0,012	0,244	0,120	0,201	0,008	103	2166	21,029	0,340	2,379	178
114	0,182	0,018	0,243	0,196	0,129	0,025	0,232	0,030	0,180	0,183	0,295	0,032	0,191	0,021	0,454	0,265	0,234	0,026	0,184	0,023	0,292	0,162	0,188	0,021	62	668	10,774	0,984	3,629	94
115	0,194	0,009	0,301	0,089	0,203	0,008	0,206	0,012	0,323	0,115	0,238	0,007	0,229	0,017	0,735	0,112	0,286	0,010	0,180	0,010	0,322	0,097	0,207	0,006	230	3870	16,826	0,996	3,713	353
116	0,155	0,018	0,276	0,218	0,144	0,014	0,153	0,025	0,216	0,190	0,148	0,021	0,167	0,030	0,205	0,255	0,203	0,024	0,155	0,026	0,256	0,230	0,179	0,020	77	776	10,078	0,506	4,558	171
117	0,224	0,025	0,293	0,249	0,227	0,016	0,210	0,034	0,396	0,280	0,274	0,022	0,224	0,040	0,304	0,212	0,289	0,028	0,251	0,026	0,288	0,232	0,294	0,022	47	557	11,851	0,319	2,340	90
118	0,171	0,011	0,381	0,146	0,144	0,006	0,165	0,011	0,310	0,153	0,191	0,005	0,197	0,022	0,708	0,134	0,214	0,010	0,193	0,011	0,314	0,171	0,189	0,012	87	3821	43,920	0,885	3,805	158
119	0,049	0,017	0,224	0,279	0,039	0,007	0,156	0,041	0,191	0,271	0,055	0,023	0,173	0,034	0,175	0,266	0,075	0,020	0,185	0,043	0,187	0,278	0,055	0,020	54	565	10,463	0,426	3,685	135
120	0,226	0,008	0,369	0,110	0,264	0,009	0,215	0,012	0,304	0,138	0,282	0,011	0,220	0,016	0,426	0,158	0,305	0,014	0,221	0,008	0,259	0,109	0,265	0,011	69	3685	53,406	0,493	2,420	112
121	0,145	0,014	0,338	0,194	0,113	0,005	0,192	0,010	0,192	0,158	0,150	0,011	0,176	0,037	0,877	0,113	0,165	0,013	0,232	0,020	0,205	0,181	0,154	0,013	69	2792	40,464	0,855	3,696	123
122	0,196	0,015	0,328	0,100	0,187	0,010	0,213	0,017	0,228	0,083	0,230	0,010	0,221	0,016	0,172	0,084	0,267	0,010	0,192	0,016	0,319	0,097	0,209	0,010	107	2816	26,318	0,850	2,963	179
123	0,207	0,043	0,425	0,389	0,140	0,022	0,362	0,054	0,667	0,422	0,442	0,061	0,268	0,066	0,450	0,409	0,298	0,055	0,321	0,058	0,383	0,384	0,194	0,038	34	493	14,500	0,353	2,794	64
124	0,117	0,022	0,283	0,307	0,059	0,013	0,284	0,049	0,338	0,269	0,225	0,044	0,192	0,040	0,336	0,382	0,102	0,022	0,196	0,024	0,198	0,273	0,060	0,011	22	495	22,500	0,591	2,091	30
125	0,220	0,018	0,346	0,173	0,232	0,010	0,312	0,020	0,481	0,258	0,361	0,023	0,371	0,027	0,979	0,047	0,339	0,028	0,219	0,020	0,358	0,208	0,300	0,017	75	913	12,173	0,973	3,227	88
126	0,126	0,026	0,358	0,302	0,117	0,014	0,252	0,039	0,224	0,288	0,168	0,036	0,233	0,042	0,212	0,249	0,125	0,025	0,141	0,036	0,358	0,320	0,152	0,027	28	546	19,500	0,357	2,250	43
127	0,295	0,037	0,353	0,425	0,204	0,031	0,426	0,047	0,406	0,408	0,457	0,072	0,355	0,057	0,264	0,362	0,348	0,055	0,320	0,055	0,478	0,422	0,193	0,034	34	420	12,353	0,382	3,324	65
128	0,761	0,103	0,750	0,433	0,022	0,007	0,812	0,160	0,750	0,433	0,147	0,105	0,898	0,142	0,750	0,433	0,053	0,029	0,188	0,038	0,750	0,433	0,803	0,105	24	592	24,667	0,167	3,333	44
129	0,325	0,104	0,750	0,433	0,200	0,029	0,535	0,124	0,750	0,433	0,433	0,153	0,565	0,096	0,750	0,433	0,208	0,050	0,445	0,102	0,750	0,433	0,200	0,029	10	172	17,200	0,000	2,300	18
130	0,172	0,022	0,345	0,315	0,221	0,016	0,269	0,041	0,261	0,270	0,264	0,030	0,275	0,035	0,352	0,354	0,260	0,029	0,220	0,032	0,321	0,302	0,238	0,024	49	743	15,163	0,327	3,265	103
131	0,043	0,034	0,385	0,284	0,011	0,001	0,039	0,034	0,296	0,313	0,014	0,002	0,040	0,034	0,404	0,289	0,013	0,002	0,038	0,032	0,248	0,293	0,018	0,003	34	4720	138,824	0,588	5,588	102
132	0,612	0,036	0,450	0,497	0,070	0,016	0,745	0,084	0,450	0,497	0,218	0,044	0,825	0,096	0,450	0,497	0,134	0,049	0,185	0,051	0,543	0,449	0,066	0,017	13	219	16,846	0,923	3,615	27
133	0,126	0,043	0,274	0,131	0,121	0,017	0,136	0,013	0,138	0,123	0,118	0,015	0,129	0,016	0,160	0,102	0,125	0,017	0,145	0,013	0,126	0,121	0,114	0,017	217	8512	39,226	0,677	2,194	288
134	0,028	0,013	0,280	0,320	0,017	0,004	0,039	0,006	0,201	0,262	0,030	0,009	0,043	0,007	0,208	0,261	0,025	0,006	0,040	0,009	0,126	0,144	0,025	0,004	42	3219	76,643	0,476	2,405	68
135	0,184	0,025	0,199	0,319	0,090	0,012	0,214	0,048	0,253	0,333	0,171	0,046	0,351	0,077	0,174	0,272	0,201	0,051	0,231	0,044	0,212	0,316	0,096	0,027	33	402	12,182	0,545	3,576	80
136	0,275	0,015	0,266	0,188	0,201	0,013	0,294	0,024	0,294	0,250	0,320	0,022	0,298	0,036	0,715	0,207	0,340	0,027	0,241	0,020	0,372	0,257	0,230	0,016	45	908	20,178	0,822	2,356	48
137	0,174	0,026	0,236	0,247	0,105	0,012	0,189	0,038	0,286	0,294	0,129	0,018	0,141	0,042	0,360	0,306	0,199	0,024	0,127	0,025	0,289	0,344	0,411	0,041	32	389	12,156	0,438	3,969	107
138	0,255	0,012	0,364	0,111	0,276	0,008	0,263	0,012	0,261	0,109	0,320	0,011	0,249	0,015	0,479	0,099	0,369	0,012	0,231	0,011	0,248	0,098	0,290	0,011	166	4718	28,422	0,849	3,012	279
139	0,282	0,009	0,357	0,108	0,258	0,013	0,237	0,014	0,379	0,134	0,278	0,012	0,257	0,010	0,281	0,123	0,306	0,010	0,248	0,010	0,377	0,130	0,255	0,011	113	4716	41,735	0,805	1,965	175
140	0,250	0,021	0,284	0,170	0,226	0,014	0,235	0,025	0,240	0,163	0,260	0,024	0,224	0,022	0,265	0,173	0,264	0,023	0,220	0,027	0,362	0,165	0,235	0,014	53	800	15,094	0,472	2,113	82
141	0,196	0,010	0,303	0,087	0,188	0,007	0,209	0,010	0,226	0,062	0,228	0,008	0,192	0,010	0,409	0,077	0,245	0,011	0,171	0,010	0,318	0,091	0,206	0,008	237	3688	15,561	0,422	2,447	308
142	0,184	0,007	0,365	0,113	0,145	0,005	0,200	0,011	0,266	0,105	0,180	0,008	0,165	0,010	0,415	0,104	0,186	0,008	0,204	0,010	0,260	0,120	0,163	0,006	64	9239	144,359	1,000	3,813	105
143	0,221	0,024	0,546	0,408	0,155	0,032	0,636	0,083	0,250	0,433	0,333	0,059	0,233	0,052	0,525	0,425	0,222	0,048	0,261	0,043	0,427	0,448	0,178	0,048	34	690	20,294	0,147	2,559	53
144	0,160	0,012	0,325	0,171	0,133	0,010	0,216	0,031	0,194	0,222	0,153	0,013	0,224	0,034	0,275	0,166	0,191	0,015	0,176	0,025	0,314	0,228	0,129	0,012	57	1013	17,772	0,825	2,561	103
145	0,256	0,037	0,600	0,490	0,203	0,028	0,550	0,079	0,500	0,500	0,429	0,084	0,221	0,055	0,597	0,460	0,250	0,065	0,276	0,035	0,567	0,481	0,226	0,043	29	524	18,069	0,172	2,483	43
146	0,037	0,014	0,341	0,390	0,001	0,003	0,147	0,038	0,379	0,429	0,125	0,045	0,072	0,055	0,360	0,377	0,032	0,023	0,065	0,031	0,166	0,355	0,013	0,008	18	314	17,444	0,389	2,278	32
147	0,150	0,049	0,574	0,267	0,120	0,015	0,175	0,051	0,551	0,265	0,158	0,017	0,147	0,032	0,397	0,276	0,120	0,016	0,160	0,047	0,396	0,281	0,119	0,015	24	3232	134,667	0,417	1,958	34
148	0,106	0,020	0,352	0,259	0,068	0,006	0,165	0,023	0,255	0,232	0,146	0,019	0,148	0,019	0,861	0,117	0,105	0,017	0,478	0,386	0,514	0,379	0,958	0,021	54	946	17,519	0,722	5,741	103
149	0,263	0,014	0,376	0,151	0,345	0,012	0,297	0,022	0,423	0,144	0,408	0,022	0,307	0,021	0,427	0,156	0,386	0,019	0,301	0,018	0,388	0,144	0,386	0,018	97	1138	11,732	0,412	2,010	149
150	0,163	0,021	0,313	0,282	0,173	0,014	0,303	0,062	0,608	0,278	0,413	0,060	0,187	0,037	0,606	0,300	0,256	0,025	0,018	0,004	0,168	0,237	0,981	0,005	30	854	28,467	1,000	2,600	38
151	0,122	0,031	0,355	0,423	0,123	0,014	0,261	0,069	0,442	0,460	0,304	0,040	0,224	0,055	0,461	0,450	0,255	0,053	0,222	0,052	0,224	0,366	0,158	0,032	20	231	11,550	0,450	2,600	38
152	0,160	0,034	0,237	0,273	0,168	0,016	0,258	0,039	0,333	0,344	0,305	0,033	0,187	0,035	0,269	0,302	0,239	0,032	0,217	0,024	0,192	0,240	0,265	0,021	23	386	16,783	0,348	3,174	49
153	0,088	0,017	0,253	0,226	0,044	0,009	0,203	0,034	0,181	0,193	0,056	0,017	0,214	0,041	0,143	0,193	0,071	0,019	0,240	0,030	0,219	0,189	0,053	0,011	57	580	10,175	0,404	3,632	141
154	0,196	0,006	0,266	0,050	0,208	0,004	0,313	0,011	0,207	0,062	0,437	0,016	0,242	0,012	0,313	0,084	0,326	0,010	0,306	0,007	0,311	0,082	0,367	0,007	406	4614	11,365	1,000	3,352	498
155	0,285	0,008	0,372	0,070	0,303	0,007	0,266	0,009	0,288	0,082	0,338	0,009	0,277	0,013	0,295	0,087	0,409	0,011	0,256	0,006	0,333	0,074	0,334	0,009	110	5656	51,418	0,709	2,545	161
156	0,145	0,018	0,337	0,113	0,131	0,006	0,170	0,011	0,162	0,098	0,152	0,007	0,154	0,017	0,204	0,091	0,165	0,013	0,166	0,010	0,163	0,075	0,123	0,009	48	3457	72,021	0,771	2,813	87
157	0,256	0,023	0,371	0,151	0,340	0,015	0,279	0,022	0,400	0,183	0,379	0,020	0,301	0,019	0,353	0,179	0,384	0,014	0,277	0,020	0,344	0,183	0,388	0,016	99	1167	11,788	0,404	2,010	153
158	0,231	0,011	0,279	0,101	0,251	0,009	0,242	0,018	0,328	0,125	0,287	0,013	0,371	0,015	0,994	0,021	0,339	0,017	0,238	0,014	0,224	0,121	0,319	0,013	156	1864	11,949	0,987	2,776	187
159	0,302	0,011	0,435	0,073	0,392	0,008	0,295	0,012	0,452	0,096	0,425	0,011	0,263	0,013	0,331	0,085	0,449	0,015	0,268	0,011	0,357	0,086	0,407	0,010	108	5302	49,093	0,435	4,630	280
160	0,241	0,007	0,327	0,079	0,241	0,010	0,272	0,009	0,263	0,110	0,285	0,011	0,263	0,008	0,353	0,108	0,333	0,009	0,280	0,007	0,321	0,098	0,277	0,008	248	5818	23,460	0,919	3,230	357
161	0,237	0,030	0,381	0,259	0,237	0,020	0,298	0,033	0,273	0,203	0,323	0,027	0,314	0,031	0,409	0,254	0,316	0,030	0,263	0,031	0,335	0,209	0,270	0,024	40	748	18,700	0,625	1,875	44
162	0,234	0,007	0,301	0,142	0,213	0,008	0,206	0,011	0,316	0,132	0,218	0,011	0,223	0,013	0,403	0,190	0,224	0,009	0,187	0,012	0,288	0,133	0,239	0,011	70	2557	36,529	0,400	1,700	90
163	0,256	0,033	0,367	0,336	0,289	0,037	0,433	0,058	0,100	0,300	0,546	0,060	0,510	0,055	0,100	0,300	0,464	0,063	0,276	0,050	0,464	0,333	0,315	0,029	20	253	12,650	0,900	2,050	24
164	0,336	0,027	0,399	0,346	0,313	0,023	0,324	0,059	0,330	0,311	0,547	0,043	0,496	0,052	0,899	0,247	0,475	0,059	0,340	0,030	0,301	0,320	0,373	0,034	45	593	13,178	0,667	2,778	64
165	0,223	0,016	0,329	0,198	0,239	0,010	0,235	0,028	0,296	0,118	0,329	0,029	0,278	0,027	0,423	0,204	0,344	0,021	0,250	0,023	0,377	0,168	0,263	0,016	74	849	11,473	0,878	2,527	98
166	0,192	0,011	0,295	0,115	0,140	0,009	0,233	0,028	0,296	0,162	0,182	0,013	0,255	0,025	0,257	0,118	0,252	0,018	0,207	0,016	0,340	0,115	0,175	0,012	107	1417	13,243	0,692	2,682	186
167	0,150	0,040	0,225	0,339	0,147	0,023	0,533	0,078	0,442	0,472	0,406	0,066	0,329	0,075	0,342	0,449	0,188	0,036	0,228	0,048	0,175	0,331	0,111	0,031	35	408	11,657	0,400	2,743	54
168	0,083	0,055	0,267	0,203	0,062	0,020	0,150	0,101	0,699	0,205	0,161	0,054	0,112	0,099	0,675	0,196	0,112	0,024	0,014	0,005	0,064	0,108	0,991	0,003	30	1827	60,900	1,000	2,600	38
169	0,182	0,009	0,334	0,099	0,158	0,009	0,269	0,014	0,374	0,139	0,255	0,012	0,230	0,016	0,649	0,128	0,242	0,015	0,259	0,009	0,224	0,122	0,207	0,009	153	3428	22,405	0,974	3,510	166
170	0,210	0,016	0,240	0,234	0,205	0,013	0,234	0,021	0,365	0,306	0,256	0,026	0,239	0,027	0,335	0,308	0,261	0,020	0,269	0,027	0,266	0,258	0,310	0,021	60	782	13,033	0,350	2,450	118
171	0,119	0,027	0,391	0,281	0,080	0,013	0,174	0,024	0,239	0,288	0,137	0,030	0,137	0,032	0,530	0,392	0,079	0,009	0,175	0,022	0,174	0,261	0,069	0,013	24	709	29,542	0,792	3,500	45
172	0,314	0,116	0,485	0,479	0,136	0,025	0,591	0,066	0,450	0,497	0,389	0,060	0,382	0,075	0,475	0,487	0,194	0,056	0,468	0,107	0,485	0,479	0,358	0,081	30	486	16,200	0,233	3,000	50
173	0,270	0,009	0,339	0,077	0,321	0,007	0,302	0,010	0,332	0,082	0,440	0,018	0,515	0,011	0,997	0,008	0,491	0,010	0,300	0,011	0,470	0,096	0,410	0,011	266	3419	12,853	1,000	2,139	287
174	0,212	0,024	0,321	0,347	0,167	0,013	0,277	0,038	0,312	0,310	0,282	0,030	0,287	0,041	0,258	0,282	0,255	0,040	0,216	0,028	0,384	0,325	0,257	0,026	26	401	15,423	0,308	2,308	50
175	0,167	0,101	0,572	0,457	0,048	0,016	0,742	0,073	0,450	0,497	0,217	0,039	0,249	0,055	0,558	0,490	0,081	0,038	0,258	0,059	0,522	0,487	0,063	0,031	11	149	13,545	0,000	0,000	20
176	0,249	0,022	0,362	0,340	0,187	0,023	0,275	0,037	0,285	0,350	0,284	0,051	0,197	0,057	0,223	0,281	0,284	0,043	0,207	0,051	0,371	0,379	0,229	0,023	31	327	10,548	1,000	2,871	49
177	0,769	0,043	0,450	0,497	0,064	0,016	0,750	0,073	0,450	0,497	0,222	0,068	0,818	0,072	0,450	0,497	0,138	0,043	0,172	0,043	0,629	0,466	0,128	0,022	16	268	16,750	0,563	2,688	27
178	0,258	0,014	0,287	0,260	0,186	0,011	0,238	0,020	0,386	0,333	0,246	0,026	0,267	0,025	0,348	0,274	0,249	0,024	0,253	0,028	0,328	0,255	0,254	0,019	48	576	12,000	0,313	2,354	93
179	0,219	0,032	0,539	0,404	0,148	0,024	0,604	0,062	0,250	0,433	0,364	0,063	0,691	0,069	0,250	0,433	0,273	0,071	0,294	0,060	0,380	0,402	0,188	0,031	16	323	20,188	0,750	1,750	17
180	0,763	0,057	0,455	0,493	0,147	0,035	0,308	0,069	0,645	0,422	0,240	0,045	0,282	0,049	0,573	0,440	0,187	0,081	0,262	0,045	0,591	0,424	0,161	0,039	15	315	21,000	0,400	2,267	29
181	0,920	0,110	0,800	0,400	0,030	0,017	0,814	0,083	0,800	0,400	0,154	0,077	0,890	0,103	0,800	0,400	0,061	0,033	0,886	0,106	0,800	0,400	0,065	0,034	12	364	30,333	0,167	3,000	23
182	0,237	0,006	0,254	0,100	0,168	0,008	0,266	0,012	0,309	0,103	0,217	0,009	0,223	0,014	0,628	0,125	0,239	0,011	0,254	0,010	0,227	0,114	0,188	0,008	194	4556	23,485	0,969	3,201	188
183	0,225	0,010	0,351	0,149	0,253	0,011	0,221	0,016	0,380	0,192	0,286	0,013	0,212	0,014	0,391	0,223	0,289	0,016	0,206	0,018	0,284	0,144	0,259	0,008	38	2450	64,474	0,342	2,526	72
184	0,124	0,018	0,322	0,235	0,068	0,010	0,182	0,024	0,246	0,312	0,173	0,028	0,154	0,039	0,800	0,305	0,117	0,020	0,212	0,034	0,173	0,245	0,073	0,009	28	675	24,107	0,821	2,607	46
185	0,288	0,026	0,316	0,209	0,267	0,012	0,314	0,028	0,352	0,161	0,311	0,017	0,296	0,024	0,396	0,180	0,305	0,015	0,312	0,018	0,313	0,212	0,256	0,013	70	783	11,186	0,400	2,171	111
186	0,383	0,029	0,717	0,438	0,067	0,017	0,677	0,104	0,700	0,458	0,292	0,115	0,353	0,082	0,717	0,438	0,075	0,053	0,490	0,053	0,700	0,458	0,066	0,028	17	288	16,941	0,118	2,706	27
187	0,159	0,014	0,255	0,123	0,137	0,007	0,147	0,014	0,249	0,112	0,147	0,008	0,225	0,028	0,887	0,113	0,212	0,021	0,133	0,014	0,249	0,101	0,202	0,013	116	1727	14,888	0,948	4,121	268
188	0,280	0,012	0,383	0,130	0,415	0,011	0,250	0,016	0,344	0,112	0,423	0,010	0,354	0,021	0,469	0,134	0,554	0,012	0,255	0,014	0,364	0,093	0,444	0,011	201	2391	11,896	0,950	1,965	206
189	0,200	0,049	0,327	0,318	0,194	0,024	0,291	0,024	0,535	0,312	0,314	0,045	0,238	0,052	0,513	0,309	0,228	0,021	0,228	0,046	0,211	0,223	0,192	0,016	49	1436	29,306	0,694	2,673	70
190	0,143	0,014	0,310	0,188	0,107	0,019	0,280	0,034	0,458	0,216	0,337	0,040	0,151	0,019	0,328	0,217	0,139	0,018	0,183	0,019	0,135	0,153	0,110	0,017	26	1042	40,077	0,962	2,000	31
191	0,162	0,009	0,302	0,087	0,117	0,006	0,143	0,008	0,285	0,125	0,115	0,005	0,148	0,022	0,633	0,168	0,164	0,010	0,155	0,009	0,257	0,117	0,106	0,005	69	4741	68,710	0,913	3,232	110
192	0,179	0,013	0,340	0,197	0,088	0,007	0,255	0,026	0,205	0,192	0,140	0,013	0,243	0,027	0,308	0,223	0,156	0,016	0,247	0,019	0,209	0,170	0,109	0,011	51	975	19,118	0,922	2,588	89
193	0,209	0,016	0,285	0,231	0,231	0,049	0,354	0,029	0,233	0,244	0,219	0,037	0,338	0,031	0,240	0,264	0,244	0,040	0,272	0,025	0,318	0,251	0,267	0,056	58	1073	18,500	1,000	2,672	117
194	0,136	0,010	0,311	0,104	0,095	0,007	0,117	0,008	0,237	0,117	0,113	0,007	0,117	0,015	0,431	0,146	0,158	0,011	0,123	0,009	0,220	0,129	0,093	0,004	57	4155	72,895	0,877	3,105	95
195	0,194	0,007	0,285	0,072	0,168	0,008	0,172	0,009	0,304	0,080	0,203	0,008	0,223	0,012	0,272	0,076	0,245	0,008	0,223	0,009	0,240	0,085	0,207	0,008	229	4798	20,952	0,904	3,218	343
196	0,147	0,016	0,356	0,196	0,110	0,007	0,180	0,020	0,200	0,188	0,135	0,015	0,173	0,025	0,970	0,064	0,135	0,013	0,191	0,015	0,194	0,178	0,103	0,008	38	2457	64,658	0,737	3,158	66
197	0,173	0,037	0,603	0,395	0,137	0,032	0,661	0,083	0,350	0,477	0,307	0,056	0,192	0,055	0,483	0,431	0,168	0,029	0,202	0,035	0,397	0,465	0,140	0,031	23	542	23,565	0,174	2,304	34
198	0,875	0,081	0,700	0,458	0,077	0,026	0,804	0,130	0,700	0,458	0,165	0,087	0,900	0,101	0,700	0,458	0,052	0,043	0,875	0,081	0,700	0,458	0,077	0,026	13	195	15,000	0,154	2,462	24
199	0,216	0,009	0,292	0,143	0,190	0,010	0,203	0,010	0,222	0,165	0,224	0,014	0,317	0,037	0,947	0,077	0,282	0,018	0,186	0,014	0,324	0,133	0,237	0,013	64	2782	43,469	0,953	2,875	90

Table 3: The results of the 4 algorithms for each workspace in the ChatWorks dataset, including the average accuracy and the standard deviation (20 random splits). The table also lists information characterizing each workspace.